GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2003, 13:25:57 ; Search time 36.1011 Seconds
(without alignments)
39.570 Million cell updates/sec

Title: US-09-744-282-1
Perfect score: 44
Sequence: 1 FGRSFTLAS 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	44	100.0	9	17	AAR95071	Human cartilage gl
2	44	100.0	9	21	AAY69226	Peptide fragment o
3	44	100.0	13	17	AAR95075	Peptide contg. hC
4	44	100.0	13	21	AAB03439	Gp38k-derived cell
5	44	100.0	13	21	AAY69230	Peptide fragment o
6	44	100.0	13	23	AAE20224	Human cartilage gl
7	44	100.0	15	19	AAW61598	Human Cartilage gp
8	44	100.0	16	18	AAW45139	Human cartilage gl
9	44	100.0	105	22	AAU02074	Synthetic human ta
10	44	100.0	220	22	AAU02078	Synthetic human mu
11	44	100.0	281	22	AAU02079	Synthetic multi-ta
12	44	100.0	350	22	AAU02076	Synthetic human mu
13	44	100.0	383	21	AAB03442	Gp38k protein seqi
14	44	100.0	383	23	ABG96297	Human ovarian canc
15	44	100.0	383	24	ABU56651	Lung cancer-associ
16	44	100.0	459	22	AAU02077	Synthetic multi-ta
17	41	93.2	372	23	AAE25906	Human chitinase pr
18	41	93.2	373	19	AAW40262	Human chitinase pr
19	41	93.2	373	19	AAW40261	Human chitinase pr
20	41	93.2	373	20	AAY42427	Clone of the C-ter
21	41	93.2	373	20	AAY42428	Chitinase amino ac
22	41	93.2	373	23	AAE25905	Human chitinase-re
23	41	93.2	373	23	ABB76293	Human chitinase tr
24	41	93.2	373	23	ABB76294	Human chitinase an
25	41	93.2	383	18	AAW26751	Bovine whey protei
26	41	93.2	387	18	AAW08585	Human 39 kDa chiti
27	41	93.2	464	18	AAW31497	Human chitotriosid
28	41	93.2	466	18	AAW31498	Human chitotriosid
29	41	93.2	466	18	AAW08584	Human 50 kDa chiti
30	41	93.2	466	19	AAW40259	Human chitinase pr
31	41	93.2	466	19	AAW40260	Human chitinase pr
32	41	93.2	466	20	AAY42425	MO-218 clone of hu
33	41	93.2	466	20	AAY42426	MO-13B clone of hu
34	41	93.2	466	22	AAE00432	Human chitinase pr
35	41	93.2	466	22	AAE00433	Human chitinase pr
36	41	93.2	466	23	AAE25903	Human chitinase al
37	41	93.2	466	23	AAE25904	Human chitinase al
38	41	93.2	466	23	ABB76291	Human chitinase.
39	41	93.2	466	23	ABB76292	Human chitinase.
40	41	93.2	466	24	ABP72621	Human chitotriosid
41	38	86.4	16	18	AAW45140	Human cartilage gl
42	38	86.4	16	19	AAW61599	Human Cartilage gp
43	38	86.4	199	22	ABB56452	Human single chain
44	38	86.4	199	22	ABB56456	Human single chain
45	38	86.4	208	22	ABB56475	Human single chain

ALIGNMENTS

RESULT 1
AAR95071
ID AAR95071 standard; peptide; 9 AA.
XX
AC AAR95071;
XX
DT 17-DEC-1996 (first entry)
XX
DE Human cartilage glycoprotein 39 autoantigen derived peptide (1).
XX
KW Human; cartilage; glycoprotein 39; hC gp-39; autoantigen;
KW induction; T-cell; tolerance; mediated; destruction;
KW rheumatoid arthritis; disease model; diagnosis.
XX
OS Homo sapiens.
XX
PN WO9613517-A1.
XX
PD 09-MAY-1996.
XX
PF 25-OCT-1995; 95WO-EP04201.
XX
PR 07-APR-1995; 95EP-0200886.
PR 27-OCT-1994; 94EP-0203128.
XX
PA (ALKO) AKZO NOBEL NV.
XX
PI Boots AMH, Verheijden GFM;
XX
DR WPI; 1996-239447/24.
XX
PT Peptide(s) derived from human cartilage gp-39 auto:antigen - used

```
PT   to induce specific T cell tolerance, esp. for treatment of
PT   rheumatoid arthritis, also useful in producing arthritic animal
PT   models
XX
PS   Claim 1; Page 28; 37pp; English.
XX
CC   The present peptide, derived from the human cartilage glycoprotein
CC   39 (hC gp-39) autoantigen, can be used to induce specific T-cell
CC   tolerance to hC gp-39 in a patient with T-cell mediated cartilage
CC   destruction, esp. rheumatoid arthritis (RA). It may also be used to
CC   induce RA in mice, etc. (human disease models), and to diagnose the
CC   presence of activated T-cells autoreactive with hC gp-39, i.e. to
CC   identify patients suitable for tolerising treatment.
CC   Therapeutically the peptide is given in a dose of 0.1-100 microg/kg
CC   by injection, or 10-50 microg/kg to produce a disease model. The
CC   peptide has a specific tolerising effect on autoreactive T-cells,
CC   but leaves other components of the immune system intact, and
CC   therefore should have few side effects. In an assay to determine
CC   the peptide's binding to HLA-DR mols., specifically DR4Dw4 (known
CC   to be associated with an increased risk of RA development), the
CC   peptide gave a DOT blot IC(50) (micromolar) value of 0.04 compared
CC   to 0.55 for the control peptide IHA(307-319)F.
XX
SQ   Sequence   9 AA;

  Query Match             100.0%;  Score 44;  DB 17;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 9.3e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FGRSFTLAS 9
              |||||||||
Db          1 FGRSFTLAS 9


RESULT 2
AAY69226
ID   AAY69226 standard; peptide; 9 AA.
XX
AC   AAY69226;
XX
DT   30-MAY-2000  (first entry)
XX
DE   Peptide fragment of human cartilage gp-39 (HC gp-39) protein.
XX
KW   Human cartilage gp-39 protein; HC gp-39 protein; T-cell activator;
KW   inflammatory disease; lymphocyte; antigen; bystander suppression;
KW   antigen-specific T-cell tolerance; autoimmune disorder;
KW   inflammatory disease; rheumatoid arthritis;
KW   autoimmune inflammatory disease; autoantigen.
XX
OS   Homo sapiens.
XX
PN   WO200004917-A2.
XX
PD   03-FEB-2000.
XX
PF   19-JUL-1999;   99WO-EP05331.
XX
PR   23-JUL-1998;   98EP-0202471.
XX
PA   (ALKU ) AKZO NOBEL NV.
XX
PI   Miltenburg AMM,  Boots AMH;
XX
DR   WPI; 2000-182537/16.
XX
PT   Use of cartilage protein or peptides for treating inflammatory diseases
PT   especially rheumatoid arthritis -
XX
PS   Claim 2; Page 17; 29pp; English.
XX
CC   AAY69226-33 represent fragments of the human cartilage gp-39 (HC gp-39)
CC   protein. The peptides are T-cell activators. The peptides are used for
CC   preventing inflammatory diseases, by modulating lymphocytes reactive
CC   to antigens other than HC gp-39, present in the same tissue of
CC   HC gp-39. By the induction of antigen-specific T-cell tolerance,
CC   autoimmune disorders can be treated by bystander suppression. The
CC   HC gp-39 peptides are useful for preventing inflammatory diseases
CC   especially rheumatoid arthritis. The fragments may also be used for
CC   treating autoimmune inflammatory diseases. The HC gp-39 peptides are
CC   especially beneficial in autoimmune conditions in which the
CC   autoantigens associated with the condition are unknown.
XX
SQ   Sequence   9 AA;

  Query Match             100.0%;  Score 44;  DB 21;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 9.3e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FGRSFTLAS 9
              |||||||||
Db          1 FGRSFTLAS 9


RESULT 3
AAR95075
ID   AAR95075 standard; peptide; 13 AA.
XX
AC   AAR95075;
XX
DT   17-DEC-1996  (first entry)
XX
DE   Peptide contg. hC gp-39 autoantigen derived peptide (I).
XX
KW   Human; cartilage; glycoprotein 39; hC gp-39; autoantigen;
KW   induction; T-cell; tolerance; mediated; destruction;
KW   rheumatoid arthritis; disease model; diagnosis.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         3..11
FT                   /note= "HC gp-39 peptide (I)"
XX
PN   WO9613517-A1.
XX
PD   09-MAY-1996.
XX
PF   25-OCT-1995;   95WO-EP04201.
XX
PR   07-APR-1995;   95EP-0200886.
PR   27-OCT-1994;   94EP-0203128.
XX
PA   (ALKU ) AKZO NOBEL NV.
XX
PI   Boots AMH,  Verheijden GFM;
XX
DR   WPI; 1996-239447/24.
XX
PT   Peptide(s) derived from human cartilage gp-39 auto:antigen  - used
PT   to induce specific T cell tolerance, esp. for treatment of
PT   rheumatoid arthritis, also useful in producing arthritic animal
PT   models
XX
PS   Claim 2; Page 28; 37pp; English.
XX
CC   The present peptide, contg. a peptide derived from the human
CC   cartilage glycoprotein 39 (hC gp-39) autoantigen, can be used to
CC   induce specific T-cell tolerance to hC gp-39 in a patient with
CC   T-cell mediated cartilage destruction, esp. rheumatoid arthritis
CC   (RA). It may also be used to induce RA in mice, etc. (human disease
CC   models), and to diagnose the presence of activated T-cells
CC   autoreactive with hC gp-39, i.e. to identify patients suitable for
CC   tolerising treatment.
CC   Therapeutically the peptide is given in a dose of 0.1-100 microg/kg
CC   by injection, or 10-50 microg/kg to produce a disease model. The
```

```
CC   peptide has a specific tolerising effect on autoreactive T-cells,
CC   but leaves other components of the immune system intact, and
CC   therefore should have few side effects.
XX
SQ   Sequence   13 AA;

  Query Match           100.0%;  Score 44;  DB 17;  Length 13;
  Best Local Similarity  100.0%;  Pred. No. 0.024;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FGRSFTLAS 9
              |||||||||
Db          3 FGRSFTLAS 11


RESULT 4
AAB03439
ID   AAB03439 standard; Peptide; 13 AA.
XX
AC   AAB03439;
XX
DT   03-JAN-2001  (first entry)
XX
DE   Gp38k-derived cell migration peptide #5.
XX
KW   Gp38k; chemoattractant; cell migration; wound healing; angiogenesis;
KW   cancer; vascular trauma; vascular disease; atherosclerosis; restenosis.
XX
OS   Unidentified.
XX
PN   WO200034469-A1.
XX
PD   15-JUN-2000.
XX
PF   10-DEC-1999;   99WO-US29262.
XX
PR   11-DEC-1998;   98US-0111856.
XX
PA   (UYNY ) UNIV NEW YORK STATE RES FOUND.
XX
PI   Millis AJT;
XX
DR   WPI; 2000-431300/37.
XX
PT   Clusterin and gp38K-related peptide capable of altering cell migration
PT   useful for treating atherosclerosis, cancer and stenosis following
PT   vascular trauma or disease -
XX
PS   Claim 1; Page 26; 43pp; English.
XX
CC   The present sequence is a peptide derived from gp38k. Gp38k is
CC   essential for the migration of vascular smooth muscle cells (VSMC). The
CC   gene and protein can, therefore, be used to promote wound healing,
CC   angiogenesis and vasculogenesis, in the treatment of stenosis following
CC   vascular trauma or disease and to treat atherosclerosis, and antisense
CC   sequences can be used to treat cancer, as angiogenesis is vital for
CC   tumour survival. This peptide can be used to alter cell migration
CC   patterns.
XX
SQ   Sequence   13 AA;

  Query Match           100.0%;  Score 44;  DB 21;  Length 13;
  Best Local Similarity  100.0%;  Pred. No. 0.024;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FGRSFTLAS 9
              |||||||||
Db          1 FGRSFTLAS 9


RESULT 5
AAY69230
ID   AAY69230 standard; peptide; 13 AA.
XX
AC   AAY69230;
XX
DT   30-MAY-2000  (first entry)
XX
DE   Peptide fragment of human cartilage gp-39 (HC gp-39) protein.
XX
KW   Human cartilage gp-39 protein; HC gp-39 protein; T-cell activator;
KW   inflammatory disease; lymphocyte; antigen; bystander suppression;
KW   antigen-specific T-cell tolerance; autoimmune disorder;
KW   inflammatory disease; rheumatoid arthritis;
KW   autoimmune inflammatory disease; autoantigen.
XX
OS   Homo sapiens.
XX
PN   WO200004917-A2.
XX
PD   03-FEB-2000.
XX
PF   19-JUL-1999;   99WO-EP05331.
XX
PR   23-JUL-1998;   98EP-0202471.
XX
PA   (ALKU ) AKZO NOBEL NV.
XX
PI   Miltenburg AMM,  Boots AMH;
XX
DR   WPI; 2000-182537/16.
XX
PT   Use of cartilage protein or peptides for treating inflammatory diseases
PT   especially rheumatoid arthritis -
XX
PS   Claim 2; Page 17; 29pp; English.
XX
CC   AAY69225-33 represent fragments of the human cartilage gp-39 (HC gp-39)
CC   protein. The peptides are T-cell activators. The peptides are used for
CC   preventing inflammatory diseases, by modulating lymphocytes reactive
CC   to antigens other than HC gp-39, present in the same tissue of
CC   HC gp-39. By the induction of antigen-specific T-cell tolerance,
CC   autoimmune disorders can be treated by bystander suppression. The
CC   HC gp-39 peptides are useful for preventing inflammatory diseases
CC   especially rheumatoid arthritis. The fragments may also be used for
CC   treating autoimmune inflammatory diseases. The HC gp-39 peptides are
CC   especially beneficial in autoimmune conditions in which the
CC   autoantigens associated with the condition are unknown.
XX
SQ   Sequence   13 AA;

  Query Match           100.0%;  Score 44;  DB 21;  Length 13;
  Best Local Similarity  100.0%;  Pred. No. 0.024;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FGRSFTLAS 9
              |||||||||
Db          3 FGRSFTLAS 11


RESULT 6
AAE20224
ID   AAE20224 standard; peptide; 13 AA.
XX
AC   AAE20224;
XX
DT   18-JUN-2002  (first entry)
XX
DE   Human cartilage glycoprotein (HC gp)-39 peptide (residues 259-271).
XX
KW   Human cartilage glycoprotein-39; HC gp-39; autoimmune disorder; MHC;
KW   major histocompatibility complex; insulin-dependent diabetes mellitus;
KW   multiple sclerosis; psoriasis; myasthenia gravis; rheumatoid arthritis;
KW   immunosuppressive; neuroprotective; therapy.
XX
OS   Homo sapiens.
```

XX
PN WO200214870-A2.
XX
PD 21-FEB-2002.
XX
PF 08-AUG-2001; 2001WO-EP09136.
XX
PR 14-AUG-2000; 2000EP-0202844.
XX
PA (ALKU) AKZO NOBEL NV.
XX
PI Steenbakkers PGA;
XX
DR WPI; 2002-269214/31.
XX
PT Method for diagnosing an autoimmune disease activity by detecting the
PT presence of an autoimmune specific major histocompatibility complex
PT (MHC)-peptide complex in a patient using antibodies specific against
PT MHC-peptide complexes -
XX
PS Disclosure; Page 43; 43pp; English.
XX
CC The patent discloses a method for diagnosing an autoimmune disease
CC activity. The method involves detecting the presence of an autoimmune
CC specific major histocompatibility complex (MHC)-peptide complex in a
CC patient suffering from an autoimmune disease with antibodies or its
CC antigen binding domains that specifically bind to a MHC-gp39-derived
CC peptide. The method and antibodies are useful for diagnosing auto-
CC immune diseases or purifying autoimmune specific MHC-peptide complexes.
CC They are also useful in therapy or for the manufacture of pharmaceutical
CC preparation for the treatment of autoimmune disorders such as insulin-
CC dependent diabetes mellitus, multiple sclerosis, myasthenia gravis,
CC psoriasis or rheumatoid arthritis. The present sequence is human
CC cartilage glycoprotein (HC gp)-39 peptide (residues 259-271). This
CC peptide is used in the exemplification of the invention.
XX
SQ Sequence 13 AA;

Query Match 100.0%; Score 44; DB 23; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.024;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FGRSFTLAS 9
|||||||||
Db 3 FGRSFTLAS 11

RESULT 7
AAW61598
ID AAW61598 standard; protein; 15 AA.
XX
AC AAW61598;
XX
DT 27-OCT-1998 (first entry)
XX
DE Human Cartilage gp-39 epitope (ii).
XX
KW Human; cartilage gp-39; monoclonal antibody; cell surface antigen;
KW rheumatoid arthritis; T-cell receptor.
XX
OS Homo sapiens.
XX
PN EP856520-A1.
XX
PD 05-AUG-1998.
XX
PF 02-DEC-1997; 97EP-0203769.
XX
PR 27-JUN-1997; 97EP-0201972.
PR 06-DEC-1996; 96EP-0203465.
XX
PA (ALKU) AKZO NOBEL NV.
PA (ALKU) AKZO NV.
XX
PI Steenbakkers PGA;
XX
DR WPI; 1998-400866/35.
XX
PT Preparation of monoclonal antibodies against cell surface antigens
PT used as diagnostic agents and to treat rheumatoid arthritis -
PT comprises immunising, then enriching, enriching and cloning B cells
PT and then selecting and cloning hybridoma
XX
PS Examples; Page 13; 23pp; English.
XX
CC The human cartilage gp-39 peptides AAW61597-W61599 were used in the
CC preparation of a monoclonal antibody against a cell surface antigen.
CC The antibody can be used for treatment of rheumatoid arthritis. It
CC is also useful against cell surface antigens and antibodies reactive
CC with the clonotypic structure of T-cell receptors are used as diagnostic
CC reagents. The process allows antibodies to be raised against
CC antigens which are present in very low amounts or which are
CC conformationally unstable without the need to screen very large numbers
CC of hybridomas and without requiring high purity Antigen for enrichment.
XX
SQ Sequence 15 AA;

Query Match 100.0%; Score 44; DB 19; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.027;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FGRSFTLAS 9
|||||||||
Db 1 FGRSFTLAS 9

RESULT 8
AAW45139
ID AAW45139 standard; peptide; 16 AA.
XX
AC AAW45139;
XX
DT 28-APR-1998 (first entry)
XX
DE Human cartilage glycoprotein 39 derived peptide #40.
XX
KW Articular cartilage; immunosuppressive therapy; antigen; autoantigen;
KW immunological tolerance; T-cell; human cartilage glycoprotein 39;
KW HC gp-39; rheumatoid arthritis; epitope.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9740068-A1.
XX
PD 30-OCT-1997.
XX
PF 22-APR-1997; 97WO-EP02051.
XX
PR 24-APR-1996; 96EP-0201106.
XX
PA (ALKU) AKZO NOBEL NV.
XX
PI Boots AMH, Verheijden GFM;
XX
DR WPI; 1997-535775/49.
XX
PT Peptide suitable for use in antigen specific immunosuppressive
PT therapy - resembles or mimics epitope present on HC gp-39, so
PT inducing systemic immunological tolerance to rheumatoid arthritis
PT auto:antigen
XX
PS Disclosure; Page 16; 82pp; English.
XX
CC The present sequence represents a peptide which resembles or mimics an
CC epitope present on human cartilage glycoprotein 39 (HC gp-39), an

```
CC   autoantigen in rheumatoid arthritis. The invention relates to peptides
CC   consisting of 16-55 amino acid residues comprising at least one of the
CC   following 19 sequences: LVCYYTSWS; FLCTHIIYS; IIYSFANIS; LKTLLSVGG;
CC   FIKSVPPFL; FDGLDLAWL; LYPGRRDKQ; YDIAKISQH; LDFTSIMTY; FISIMTYDF;
CC   FRGQEDASP; YAVGYMLRL; MLRLGAPAS; LAYYEICDF; LRGATVHRT; YLKDRQLAG;
CC   LAGAMVWAL; VWALDLDDF; or LDLDDFQGS. They can be used medically in
CC   antigen specific immunosuppressive therapy, particularly the treatment
CC   of T-cell mediated destruction of articular cartilage in autoimmune
CC   diseases (e.g. rheumatoid arthritis). They can also be used to detect
CC   activated autoreactive T cells in an individual. The peptides have a
CC   specific effect on the autoreactive T cells, thus leaving the other
CC   components of the immune system intact, unlike the non-specific
CC   suppressive effect of immunosuppressive drugs, and do not cause toxic
CC   side effects. The peptides are predominantly recognised by autoreactive
CC   T cells from rheumatoid arthritis patients, but rarely by those from
CC   healthy donors.
XX
SQ   Sequence   16 AA;

  Query Match             100.0%;  Score 44;  DB 18;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 0.029;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FGRSFTLAS 9
              |||||||||
Db          6 FGRSFTLAS 14


RESULT 9
AAU02074
ID   AAU02074 standard; Protein; 105 AA.
XX
AC   AAU02074;
XX
DT   07-SEP-2001  (first entry)
XX
DE   Synthetic human target autoantigen shGLP/RA.
XX
KW   shGLP/RA; immunogenic epitope cluster; IEC;
KW   synthetic human chondrocyte glycoprotein 69/cartilage link protein;
KW   autoantigen; autoimmune disease; multiple sclerosis;
KW   rheumatoid arthritis; insulin-dependent diabetes mellitus;
KW   myasthenia gravis; uveitis; autoimmune hepatitis; thyroiditis; insulitis;
KW   orchitis; idiopathic thrombocytopenic purpura; inflammatory disease;
KW   Crohn's disease; ulcerative colitis.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 103
FT                   /label= OTHER
FT                   /note= "Other= STOP"
XX
PN   WO200131037-A2.
XX
PD   03-MAY-2001.
XX
PF   26-OCT-2000; 2000WO-IL00688.
XX
PR   27-OCT-1999;   99IL-0132611.
XX
PA   (YEDA ) YEDA RES & DEV CO LTD.
XX
PI   Ben-Nun A,  Kerlero De Rosbo N,  Sappler GP;
XX
DR   WPI; 2001-300515/31.
DR   N-PSDB; AAS04827.
XX
PT   Novel synthetic human target autoantigen gene useful for treating
PT   autoimmune diseases such as multiple sclerosis, insulin-dependent
PT   diabetes mellitus, rheumatoid arthritis, myasthenia gravis, and uveitis
PT   -
XX
PS   Claim 173; Fig 51; 182pp; English.
XX
CC   The sequence represents shGLP/RA (synthetic human chondrocyte
CC   glycoprotein 69/cartilage link protein) containing immunogenic epitope
CC   clusters (IEC) from chondrocyte glycoprotein 69 and cartilage link
CC   protein. The synthetic human target autoantigen genes of the invention
CC   comprise sequences coding for at least 2 IECs of autoantigen(s) related
CC   to a specific autoimmune disease. The synthetic human target autoantigen
CC   genes are useful for treating autoimmune diseases such as multiple
CC   sclerosis, insulin-dependent diabetes mellitus, rheumatoid arthritis,
CC   myasthenia gravis, uveitis, autoimmune hepatitis, thyroiditis, insulitis,
CC   orchitis, idiopathic thrombocytopenic purpura, and inflammatory diseases
CC   (Crohn's disease, ulcerative colitis). The synthetic human target
CC   autoantigen genes are also useful for diagnosis and/or monitoring the
CC   progression of the autoimmune disease.
XX
SQ   Sequence   105 AA;

  Query Match             100.0%;  Score 44;  DB 22;  Length 105;
  Best Local Similarity   100.0%;  Pred. No. 0.21;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FGRSFTLAS 9
              |||||||||
Db         27 FGRSFTLAS 35


RESULT 10
AAU02078
ID   AAU02078 standard; Protein; 220 AA.
XX
AC   AAU02078;
XX
DT   07-SEP-2001  (first entry)
XX
DE   Synthetic human multi-target autoantigen Y-RAPc.
XX
KW   Y-RAPc; immunogenic epitope cluster; IEC;
KW   autoantigen; shCollagen/RA; shAggrecan/RA; shGLP/RA;
KW   autoimmune disease; multiple sclerosis; rheumatoid arthritis;
KW   insulin-dependent diabetes mellitus; myasthenia gravis; uveitis;
KW   autoimmune hepatitis; thyroiditis; insulitis; orchitis;
KW   idiopathic thrombocytopenic purpura; inflammatory disease;
KW   Crohn's disease; ulcerative colitis; shMultiTAG.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 218
FT                   /label= OTHER
FT                   /note= "Other= STOP"
XX
PN   WO200131037-A2.
XX
PD   03-MAY-2001.
XX
PF   26-OCT-2000; 2000WO-IL00688.
XX
PR   27-OCT-1999;   99IL-0132611.
XX
PA   (YEDA ) YEDA RES & DEV CO LTD.
XX
PI   Ben-Nun A,  Kerlero De Rosbo N,  Sappler GP;
XX
DR   WPI; 2001-300515/31.
DR   N-PSDB; AAS04839.
XX
PT   Novel synthetic human target autoantigen gene useful for treating
PT   autoimmune diseases such as multiple sclerosis, insulin-dependent
PT   diabetes mellitus, rheumatoid arthritis, myasthenia gravis, and uveitis
PT   -
```

```
XX
PS   Claim 185; Fig 59; 182pp; English.
XX
CC   The sequence represents synthetic human multi-target autoantigen
CC   (shMultiTAG) Y-RAPa consisting of regions encoding preferred epitopes
CC   from shCollagen/RA, shAggrecan/RA and shGLP/RA. The synthetic human
CC   target autoantigen genes of the invention comprise sequences coding for
CC   at least 2 IECs of autoantigen(s) related to a specific autoimmune
CC   disease. The synthetic human target autoantigen genes are useful for
CC   treating autoimmune diseases such as multiple sclerosis, insulin-
CC   dependent diabetes mellitus, rheumatoid arthritis, myasthenia gravis,
CC   uveitis, autoimmune hepatitis, thyroiditis, insulitis, orchitis,
CC   idiopathic thrombocytopenic purpura, and inflammatory diseases (Crohn's
CC   disease, ulcerative colitis). The synthetic human target autoantigen
CC   genes are also useful for diagnosis and/or monitoring the progression of
CC   the autoimmune disease.
XX
SQ   Sequence   220 AA;

  Query Match             100.0%;  Score 44;  DB 22;  Length 220;
  Best Local Similarity   100.0%;  Pred. No. 0.44;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FGRSFTLAS 9
              |||||||||
Db        183 FGRSFTLAS 191


RESULT 11
AAU02079
ID   AAU02079 standard; Protein; 281 AA.
XX
AC   AAU02079;
XX
DT   07-SEP-2001  (first entry)
XX
DE   Synthetic multi-target autoantigen Y-RAPd.
XX
KW   Y-RAPd; immunogenic epitope cluster; IEC;
KW   autoantigen; shCollagen/RA; shAggrecan/RA; shGLP/RA; shHSP/RA;
KW   autoimmune disease; multiple sclerosis; rheumatoid arthritis;
KW   insulin-dependent diabetes mellitus; myasthenia gravis; uveitis;
KW   autoimmune hepatitis; thyroiditis; insulitis; orchitis;
KW   idiopathic thrombocytopenic purpura; inflammatory disease;
KW   Crohn's disease; ulcerative colitis; shMultiTAG.
XX
OS   Synthetic.
OS   Homo sapiens.
OS   Escherichia coli.
OS   Mycobacterium tuberculosis.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 279
FT                   /label= OTHER
FT                   /note= "Other= STOP"
XX
PN   WO200131037-A2.
XX
PD   03-MAY-2001.
XX
PF   26-OCT-2000; 2000WO-IL00688.
XX
PR   27-OCT-1999;   99IL-0132611.
XX
PA   (YEDA ) YEDA RES & DEV CO LTD.
XX
PI   Ben-Nun A,  Keriero De Rosbo N,  Sappler GP;
XX
DR   WPI; 2001-300515/31.
DR   N-PSDB; AAS04840.
XX
PT   Novel synthetic human target autoantigen gene useful for treating
PT   autoimmune diseases such as multiple sclerosis, insulin-dependent
PT   diabetes mellitus, rheumatoid arthritis, myasthenia gravis, and uveitis
PT   -
XX
PS   Claim 188; Fig 61; 182pp; English.
XX
CC   The sequence represents synthetic human multi-target autoantigen
CC   (shMultiTAG) Y-RAPd consisting of preferred epitopes encoded by
CC   shCollagen/RA, shAggrecan/RA, shGLP/RA and shHSP/RA. The synthetic human
CC   target autoantigen genes of the invention comprise sequences coding for
CC   at least 2 IECs of autoantigen(s) related to a specific autoimmune
CC   disease. The synthetic human target autoantigen genes are useful for
CC   treating autoimmune diseases such as multiple sclerosis, insulin-
CC   dependent diabetes mellitus, rheumatoid arthritis, myasthenia gravis,
CC   uveitis, autoimmune hepatitis, thyroiditis, insulitis, orchitis,
CC   idiopathic thrombocytopenic purpura, and inflammatory diseases (Crohn's
CC   disease, ulcerative colitis). The synthetic human target autoantigen
CC   genes are also useful for diagnosis and/or monitoring the progression of
CC   the autoimmune disease.
XX
SQ   Sequence   281 AA;

  Query Match             100.0%;  Score 44;  DB 22;  Length 281;
  Best Local Similarity   100.0%;  Pred. No. 0.57;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FGRSFTLAS 9
              |||||||||
Db        183 FGRSFTLAS 191


RESULT 12
AAU02076
ID   AAU02076 standard; Protein; 350 AA.
XX
AC   AAU02076;
XX
DT   07-SEP-2001  (first entry)
XX
DE   Synthetic human multi-target autoantigen Y-RAPa.
XX
KW   Y-RAPa; immunogenic epitope cluster; IEC;
KW   autoantigen; shCollagen/RA; shAggrecan/RA; shGLP/RA;
KW   autoimmune disease; multiple sclerosis; rheumatoid arthritis;
KW   insulin-dependent diabetes mellitus; myasthenia gravis; uveitis;
KW   autoimmune hepatitis; thyroiditis; insulitis; orchitis;
KW   idiopathic thrombocytopenic purpura; inflammatory disease;
KW   Crohn's disease; ulcerative colitis; shMultiTAG.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 348
FT                   /label= OTHER
FT                   /note= "Other= STOP"
XX
PN   WO200131037-A2.
XX
PD   03-MAY-2001.
XX
PF   26-OCT-2000; 2000WO-IL00688.
XX
PR   27-OCT-1999;   99IL-0132611.
XX
PA   (YEDA ) YEDA RES & DEV CO LTD.
XX
PI   Ben-Nun A,  Kerlero De Rosbo N,  Sappler GP;
XX
DR   WPI; 2001-300515/31.
DR   N-PSDB; AAS04837.
XX
PT   Novel synthetic human target autoantigen gene useful for treating
PT   autoimmune diseases such as multiple sclerosis, insulin-dependent
```

```
PT   diabetes mellitus, rheumatoid arthritis, myasthenia gravis, and uveitis
PT   -
XX
PS   Claim 179; Fig 55; 182pp; English.
XX
CC   The sequence represents synthetic human multi-target autoantigen
CC   (shMultiTAG) Y-RAPa consisting of regions encoding shCollagen/RA,
CC   shAggrecan/RA and shGLP/RA. The synthetic human target autoantigen genes
CC   of the invention comprise sequences coding for at least 2 IECs of
CC   autoantigen(s) related to a specific autoimmune disease. The synthetic
CC   human target autoantigen genes are useful for treating autoimmune
CC   diseases such as multiple sclerosis, insulin-dependent diabetes
CC   mellitus, rheumatoid arthritis, myasthenia gravis, uveitis, autoimmune
CC   hepatitis, thyroiditis, insulitis, orchitis, idiopathic thrombocytopenic
CC   purpura, and inflammatory diseases (Crohn's disease, ulcerative colitis).
CC   The synthetic human target autoantigen genes are also useful for
CC   diagnosis and/or monitoring the progression of the autoimmune disease.
XX
SQ   Sequence   350 AA;

  Query Match             100.0%;  Score 44;  DB 22;  Length 350;
  Best Local Similarity   100.0%;  Pred. No. 0.72;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FGRSFTLAS 9
              |||||||||
Db        272 FGRSFTLAS 280


RESULT 13
AAB03442
ID   AAB03442 standard; Protein; 383 AA.
XX
AC   AAB03442;
XX
DT   03-JAN-2001  (first entry)
XX
DE   Gp38k protein sequence.
XX
KW   Gp38k; chemoattractant; cell migration; wound healing; angiogenesis;
KW   cancer; vascular trauma; vascular disease; atherosclerosis; restenosis;
XX
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..24
FT                   /label= signal_peptide
FT   Region          30..37
FT                   /note= "peptide antibody"
FT   Modified-site   59..61
FT                   /label= glycosylation_site
FT   Binding-site    68..75
FT                   /label= leucine_zipper
FT   Binding-site    82..89
FT                   /label= leucine_zipper
FT   Active-site     131..136
FT                   /label= chitanase_active_site
FT   Binding-site    143..146
FT                   /label= heparin_binding_site
FT   Binding-site    147..154
FT                   /label= hyaluronic_acid_binding_site
FT   Binding-site    262..270
FT                   /label= hyaluronic_acid_binding_site
FT   Binding-site    278..281
FT                   /label= glycosaminoglycan_binding_site
FT   Region          354..357
FT                   /note= "acidic region"
FT   Binding-site    368..376
FT                   /label= hyaluronic_acid_binding_site
XX
PN   WO200034469-A1.
XX
PD   15-JUN-2000.
XX
PF   10-DEC-1999;   99WO-US29262.
XX
PR   11-DEC-1998;   98US-0111856.
XX
PA   (JYNY ) UNIV NEW YORK STATE RES FOUND.
XX
PI   Millis AJT;
XX
DR   WPI; 2000-431300/37.
XX
PT   Clusterin and gp38K-related peptide capable of altering cell migration
PT   useful for treating atherosclerosis, cancer and stenosis following
PT   vascular trauma or disease  -
XX
PS   Disclosure; Fig 2; 43pp; English.
XX
CC   The present sequence is the protein sequence of gp38k. Gp38k, a
CC   chemoattractant, is essential for the migration of vascular smooth muscle
CC   cells (VSMC). The gene and protein can, therefore, be used to promote
CC   wound healing, angiogenesis and vasculogenesis, in the treatment of
CC   stenosis following vascular trauma or disease and to treat
CC   atherosclerosis, and antisense sequences can be used to treat cancer, as
CC   angiogenesis is vital for tumour survival.
XX
SQ   Sequence   383 AA;

  Query Match             100.0%;  Score 44;  DB 21;  Length 383;
  Best Local Similarity   100.0%;  Pred. No. 0.79;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FGRSFTLAS 9
              |||||||||
Db        261 FGRSFTLAS 269


RESULT 14
ABG96297
ID   ABG96297 standard; Protein; 383 AA.
XX
AC   ABG96297;
XX
DT   11-DEC-2002  (first entry)
XX
DE   Human ovarian cancer marker OV6.
XX
KW   Human; ovarian cancer; marker; cancer; familial history; brain disorder;
KW   central nervous system disorder; bacterial meningitis; viral meningitis;
KW   Alzheimer's disease; Parkinson's disease; cerebral oedema; hydrocephalus;
KW   brain herniation; inflammation; encephalitis; testicular disorder;
KW   nontuberculous granulomatous orchitis; connective tissue disorder;
KW   heart disorder; ischaemic heart disease; atherosclerosis; neoplasm;
KW   histological type; carcinogenic; ovarian cancer marker.
XX
OS   Homo sapiens.
XX
PN   WO200271928-A2.
XX
PD   19-SEP-2002.
XX
PF   14-MAR-2002; 2002WO-US07826.
XX
PR   14-MAR-2001; 2001US-276025P.
PR   14-MAR-2001; 2001US-276026P.
PR   10-AUG-2001; 2001US-311732P.
PR   19-SEP-2001; 2001US-323580P.
PR   26-SEP-2001; 2001US-324967P.
PR   26-SEP-2001; 2001US-325102P.
PR   26-SEP-2001; 2001US-325149P.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   Monahan JE,  Gannavarapu M,  Hoersch S,  Kamatkar S,  Kovatis SG;
```

PI Meyers RE, Morrisey MP, Olandt PJ, Sen A, Vieby PO, Mills GB;
PI Bast RC, Lu K, Schmandt RE, Zhao X, Glatt K;
XX
DR WPI; 2002-723277/78.
DR N-PSDB; ABS76389.
XX
PT Assessing whether a patient is afflicted with ovarian cancer, useful in
PT assessing the stage or progression of the disease, comprises comparing
PI the expression level of a cancer marker in a sample from a patient and
PT from a non cancer patient -
XX
PS Disclosure; Page 164-165; 481pp; English.
XX
CC The present invention relates to a new method for assessing whether a
CC patient is afflicted with ovarian cancer. The method involves comparing
CC the expression level of a marker in a patient sample and the normal level
CC of expression of the marker in a control non-ovarian cancer sample, where
CC the marker is selected from 363 cancer markers described in the
CC specification. The method of the invention is useful in diagnosing or
CC characterising cancer, in detecting the presence of cancer as early as
CC possible, and the recurrence of ovarian cancer. The method may also be of
CC particular use with patients having an enhanced risk of developing
CC ovarian cancer (e.g. patients having a familial history of ovarian
CC cancer). The cancer markers may be used in the management and treatment
CC of e.g. brain and central nervous system disorders (e.g. bacterial and
CC viral meningitis, Alzheimer's disease or Parkinson's disease), brain
CC disorders (e.g. cerebral oedema, hydrocephalus or brain herniations),
CC inflammations (e.g. bacterial or viral meningitis or encephalitis),
CC testicular disorders (e.g. nontuberculous granulomatous orchitis),
CC connective tissue disorders, or heart disorders (e.g. ischaemic heart
CC disease or atherosclerosis). The compositions and methods may also be
CC used in assessing the histological type of neoplasm associated with
CC ovarian cancer, monitoring the progression of ovarian cancer,
CC determining whether ovarian cancer has metastasized or is likely to
CC metastasize, selecting a composition for inhibiting ovarian cancer,
CC assessing the ovarian carcinogenic potential of a compound, or
CC inhibiting ovarian cancer or at risk of developing ovarian cancer. The
CC present amino acid sequence represents one of the ovarian cancer markers
CC described in the invention.
XX
SQ Sequence 383 AA;

Query Match 100.0%; Score 44; DB 23; Length 383;
Best Local Similarity 100.0%; Pred. No. 0.79;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FGRSFTLAS 9
|||||||||
Db 261 FGRSFTLAS 269

RESULT 15
ABU56651
ID ABU56651 standard; Protein; 383 AA.
XX
AC ABU56651;
XX
DT 02-APR-2003 (first entry)
XX
DE Lung cancer-associated polypeptide #244.
XX
KW Lung cancer-associated polypeptide; cytostatic; emphysema;
KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
XX
OS Unidentified.
XX
PN WO200286443-A2.
XX
PD 31-OCT-2002.
XX
PF 18-APR-2002; 2002WO-US12476.
XX
PR 18-APR-2001; 2001US-284770P.
PR 10-MAY-2001; 2001US-290492P.
PR 09-NOV-2001; 2001US-339245P.
PR 13-NOV-2001; 2001US-350666P.
PR 29-NOV-2001; 2001US-334370P.
PR 12-APR-2002; 2002US-372246P.
XX
PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI Aziz N, Murray R;
XX
DR WPI; 2003-093161/08.
DR N-PSDB; ABX76380.
XX
PT Detecting a lung cancer-associated transcript in a cell from a patient
PT for treating lung cancer, by contacting a biological sample from the
PT patient with a polynucleotide that exhibits increased or decreased
PT expression in lung cancer -
XX
PS Claim 27; Page 377; 453pp; English.
XX
CC The invention relates to a method for detecting a lung cancer-associated
CC transcript in a cell from a patient, comprising contacting a biological
CC sample from the patient with a polynucleotide that selectively hybridises
CC to a sequence that is at least 80 % identical to a gene that exhibits
CC increased or decreased expression in lung cancer samples. Lung
CC cancer-associated polynucleotides and polypeptides are used for
CC identifying a compound that modulates a lung cancer-associated
CC polypeptide, for inhibiting proliferation of a lung cancer-associated
CC cell to treat lung cancer in a patient and for treating a mammal having
CC lung cancer by administering a modulatory compound identified. The
CC methods are useful for treating lung cancer, such as small cell lung
CC cancer, non-small cell lung cancer or other benign or precancerous
CC lesions, e.g. atelectasis, emphysema, bronchitis, chronic obstructive
CC pulmonary disease, fibrosis, hypersensitivity pneumonitis, interstitial
CC pulmonary fibrosis, asthma and bronchiectasis. The genes, polynucleotides
CC and polypeptides are useful for diagnostic purposes and as targets for
CC screening for therapeutic compounds that modulate lung cancer, such as
CC antibodies. Sequences ABU56408-ABU56745 represent lung cancer-associated
CC polypeptides of the invention.
XX
SQ Sequence 383 AA;

Query Match 100.0%; Score 44; DB 24; Length 383;
Best Local Similarity 100.0%; Pred. No. 0.79;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FGRSFTLAS 9
|||||||||
Db 261 FGRSFTLAS 269

Search completed: September 24, 2003, 13:41:19
Job time : 37.2261 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 24, 2003, 13:35:23 ; Search time 10.9213 Seconds
(without alignments)
34.867 Million cell updates/sec

Title: US-09-744-282-1
Perfect score: 44
Sequence: 1 FGRSFTLAS 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	44	100.0	9	1	US-08-619-645-1	Sequence 1, Appli
2	44	100.0	9	2	US-08-634-493-1	Sequence 1, Appli
3	44	100.0	13	1	US-08-619-645-5	Sequence 5, Appli
4	44	100.0	13	2	US-08-634-493-5	Sequence 5, Appli
5	44	100.0	13	4	US-09-459-749D-5	Sequence 5, Appli
6	44	100.0	15	4	US-09-405-745-2	Sequence 2, Appli
7	44	100.0	16	3	US-09-171-705-40	Sequence 40, Appl
8	44	100.0	383	4	US-09-459-749D-17	Sequence 17, Appl
9	41	93.2	373	3	US-09-039-198A-14	Sequence 14, Appl
10	41	93.2	373	3	US-09-039-198A-15	Sequence 15, Appl
11	41	93.2	373	4	US-08-877-599-14	Sequence 14, Appl
12	41	93.2	373	4	US-08-877-599-15	Sequence 15, Appl
13	41	93.2	373	4	US-09-267-574-14	Sequence 14, Appl
14	41	93.2	373	4	US-09-267-574-15	Sequence 15, Appl
15	41	93.2	387	2	US-08-486-839-6	Sequence 6, Appli
16	41	93.2	387	3	US-09-151-011-6	Sequence 6, Appli
17	41	93.2	387	4	US-09-343-623-6	Sequence 6, Appli
18	41	93.2	466	2	US-08-486-839-4	Sequence 4, Appli
19	41	93.2	466	3	US-09-151-011-4	Sequence 4, Appli
20	41	93.2	466	3	US-09-039-198A-2	Sequence 2, Appli
21	41	93.2	466	3	US-09-039-198A-4	Sequence 4, Appli
22	41	93.2	466	4	US-09-343-623-4	Sequence 4, Appli
23	41	93.2	466	4	US-08-877-599-2	Sequence 2, Appli
24	41	93.2	466	4	US-08-877-599-4	Sequence 4, Appli
25	41	93.2	466	4	US-09-267-574-2	Sequence 2, Appli
26	41	93.2	466	4	US-09-267-574-4	Sequence 4, Appli
27	38	86.4	16	3	US-09-171-705-41	Sequence 41, Appl
28	38	86.4	16	4	US-09-405-745-3	Sequence 3, Appli
29	36	81.8	93	1	US-07-612-674-12	Sequence 12, Appl
30	36	81.8	276	1	US-07-612-674-5	Sequence 5, Appli
31	36	81.8	279	1	US-07-612-674-8	Sequence 8, Appli
32	36	81.8	385	2	US-08-694-915-2	Sequence 2, Appli
33	36	81.8	416	2	US-08-694-915-4	Sequence 4, Appli
34	36	81.8	423	4	US-08-850-348A-2	Sequence 2, Appli
35	35	79.5	489	4	US-09-545-814-29	Sequence 29, Appl
36	35	79.5	554	2	US-08-524-051-2	Sequence 2, Appli
37	35	79.5	554	3	US-09-052-778-16	Sequence 16, Appl
38	35	79.5	559	4	US-09-545-814-14	Sequence 14, Appl
39	35	79.5	583	4	US-09-545-814-2	Sequence 2, Appli
40	35	79.5	583	4	US-09-545-814-5	Sequence 5, Appli
41	32	72.7	13	1	US-08-619-645-6	Sequence 6, Appli
42	32	72.7	13	2	US-08-634-493-6	Sequence 6, Appli
43	32	72.7	13	4	US-09-405-745-1	Sequence 1, Appli
44	32	72.7	408	4	US-09-107-532A-6563	Sequence 6563, Ap
45	31	70.5	168	4	US-09-325-932A-69	Sequence 69, Appl

ALIGNMENTS

```
RESULT 1
US-08-619-645-1
; Sequence 1, Application US/08619645
; Patent No. 5736507
;  GENERAL INFORMATION:
;    APPLICANT:  Boots, Anna M.H.
;    APPLICANT:  Verheijden, Gijsbertus F.M.
;    TITLE OF INVENTION:  No. 5736507el peptides derived from
;    TITLE OF INVENTION:  autoantigen for use in immunotherapy of autoimmune
;    TITLE OF INVENTION:  disease
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Akzo No. 5736507el Patent Department
;      STREET:  1300 Piccard Drive, Suite 205
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  U.S.A.
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25(EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/619,645
;      FILING DATE:  25-MAR-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Gormley, Mary E.
;      REGISTRATION NUMBER:  34,409
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 947-4433
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-619-645-1

  Query Match           100.0%;  Score 44;  DB 1;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 2.5e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FGRSFTLAS 9
              |||||||||
Db          1 FGRSFTLAS 9
```

```
RESULT 2
US-08-634-493-1
; Sequence 1, Application US/08634493
; Patent No. 5843449
;  GENERAL INFORMATION:
;    APPLICANT:  A.M.H. Boots
;    APPLICANT:  G.F.M. Verheijden
;    APPLICANT:  E.S. Bos
;    TITLE OF INVENTION:  No. 5843449el Peptides derived from autoantigen for use
;    TITLE OF INVENTION:  in Immunotherapy of Autoimmune Diseases
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Akzo No. 5843449el Patent Department
;      STREET:  1300 Piccard Drive, Suite 206
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/634,493
;      FILING DATE:  18-APR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/619,645
;      FILING DATE:  25-MAR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/EP95/04201
;      FILING DATE:  25-OCT-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  NL 942031287
;      FILING DATE:  27-OCT-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  NL 952008860
;      FILING DATE:  07-APR-1995
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 258-5200
;      TELEFAX:  (301) 977-0847
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-634-493-1

  Query Match             100.0%;  Score 44;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.5e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FGRSFTLAS 9
              |||||||||
Db          1 FGRSFTLAS 9


RESULT 3
US-08-619-645-5
; Sequence 5, Application US/08619645
; Patent No. 5736507
;  GENERAL INFORMATION:
;    APPLICANT:  Boots, Anna M.H.
;    APPLICANT:  Verheijden, Gijsbertus F.M.
;    TITLE OF INVENTION:  No. 5736507el peptides derived from
;    TITLE OF INVENTION:  autoantigen for use in immunotherapy of autoimmune
;    TITLE OF INVENTION:  disease
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Akzo No. 5736507el Patent Department
;      STREET:  1300 Piccard Drive, Suite 206
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  U.S.A.
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25(EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/619,645
;      FILING DATE:  25-MAR-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Gormley, Mary E.
;      REGISTRATION NUMBER:  34,409
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 947-4433
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  13 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-619-645-5

  Query Match             100.0%;  Score 44;  DB 1;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 9.3e-04;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FGRSFTLAS 9
              |||||||||
Db          3 FGRSFTLAS 11


RESULT 4
US-08-634-493-5
; Sequence 5, Application US/08634493
; Patent No. 5843449
;  GENERAL INFORMATION:
;    APPLICANT:  A.M.H. Boots
;    APPLICANT:  G.F.M. Verheijden
;    APPLICANT:  E.S. Bos
;    TITLE OF INVENTION:  No. 5843449el Peptides derived from autoantigen for use
;    TITLE OF INVENTION:  in Immunotherapy of Autoimmune Diseases
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Akzo No. 5843449el Patent Department
;      STREET:  1300 Piccard Drive, Suite 206
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/634,493
;      FILING DATE:  18-APR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/619,645
;      FILING DATE:  25-MAR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/EP95/04201
;      FILING DATE:  25-OCT-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  NL 942031287
;      FILING DATE:  27-OCT-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  NL 952008860
```

```
;      FILING DATE:  07-APR-1995
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 258-5200
;      TELEFAX:  (301) 977-0847
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  13 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-634-493-5

  Query Match           100.0%;  Score 44;  DB 2;  Length 13;
  Best Local Similarity  100.0%;  Pred. No. 0.0048;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FGRSFTLAS 9
              |||||||||
Db          3 FGRSFTLAS 11


RESULT 5
US-09-459-749D-5
; Sequence 5, Application US/09459749D
; Patent No. 6464975
; GENERAL INFORMATION:
;  APPLICANT: Millis, Albert J. T.
;  TITLE OF INVENTION: Compositions and Methods For Altering Cell Migration
;  FILE REFERENCE: 0794.016A
;  CURRENT APPLICATION NUMBER: US/09/459,749D
;  CURRENT FILING DATE:  1999-12-10
;  PRIOR APPLICATION NUMBER: 60/111,856
;  PRIOR FILING DATE: 1998-12-11
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
;   LENGTH: 13
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:SITE
;   NAME/KEY: SITE
;   LOCATION: (1)..(13)
;   OTHER INFORMATION: Peptide corresponding to amino acids 261- 273 of
;   OTHER INFORMATION: gp38k protein
US-09-459-749D-5

  Query Match           100.0%;  Score 44;  DB 4;  Length 13;
  Best Local Similarity  100.0%;  Pred. No. 0.0048;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FGRSFTLAS 9
              |||||||||
Db          1 FGRSFTLAS 9


RESULT 6
US-09-405-745-2
; Sequence 2, Application US/09405745
; Patent No. 6392020
;  GENERAL INFORMATION:
;    APPLICANT:  Steenbakkers PGA
;    TITLE OF INVENTION:  Method of preparing a monoclonal
;    TITLE OF INVENTION:  antibody (as amended)
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Akzo No. 6392020el Patent Dept.
;      STREET:  1300 Piccard Drive, Suite 206
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  US
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/405,745
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/985,898
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sullivan, Michael G.
;      REGISTRATION NUMBER:  35,377
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  301-948-7400
;      TELEFAX:  301-948-9751
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  15 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FRAGMENT TYPE:  internal
US-09-405-745-2

  Query Match           100.0%;  Score 44;  DB 4;  Length 15;
  Best Local Similarity  100.0%;  Pred. No. 0.0055;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FGRSFTLAS 9
              |||||||||
Db          1 FGRSFTLAS 9


RESULT 7
US-09-171-705-40
; Sequence 40, Application US/09171705
; Patent No. 6184204
; GENERAL INFORMATION:
;  APPLICANT: BOOTS, ANNA M.H.
;  APPLICANT: VERHEIJDEN, GILBERTUS F.M.
;  TITLE OF INVENTION: NOVEL PEPTIDES SUITABLE FOR USE IN ANTIGEN SPECIFIC
;  TITLE OF INVENTION: IMMUNOSUPPRESSIVE THERAPY
;  FILE REFERENCE: O/96198 US
;  CURRENT APPLICATION NUMBER: US/09/171,705
;  CURRENT FILING DATE: 1999-02-09
;  NUMBER OF SEQ ID NOS: 78
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 40
;   LENGTH: 16
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: DERIVED FROM
;   OTHER INFORMATION: SEQUENCE OF HUMAN CARTILAGE (HC) -39 PROTEIN
US-09-171-705-40

  Query Match           100.0%;  Score 44;  DB 3;  Length 16;
  Best Local Similarity  100.0%;  Pred. No. 0.006;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FGRSFTLAS 9
              |||||||||
Db          6 FGRSFTLAS 14


RESULT 8
US-09-459-749D-17
; Sequence 17, Application US/09459749D
; Patent No. 6464975
```

```
; GENERAL INFORMATION:
;  APPLICANT: Millis, Albert J. T.
;  TITLE OF INVENTION: Compositions and Methods For Altering Cell Migration
;  FILE REFERENCE: 0794.016A
;  CURRENT APPLICATION NUMBER: US/09/459,749D
;  CURRENT FILING DATE:  1999-12-10
;  PRIOR APPLICATION NUMBER: 60/111,856
;  PRIOR FILING DATE: 1998-12-11
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
;   LENGTH: 383
;   TYPE: PRT
;   ORGANISM: Sus scrofa
US-09-459-749D-17

  Query Match            100.0%;  Score 44;  DB 4;  Length 383;
  Best Local Similarity  100.0%;  Pred. No. 0.22;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FGRSFTLAS 9
              |||||||||
Db        261 FGRSFTLAS 269


RESULT 9
US-09-039-198A-14
; Sequence 14, Application US/09039198A
; Patent No. 6200951
;  GENERAL INFORMATION:
;    APPLICANT:  Gray, Patrick W.
;    APPLICANT:  Tjoelker, Larry W.
;    TITLE OF INVENTION:  CHITINASE CHITIN-BINDING FRAGMENTS
;    NUMBER OF SEQUENCES:  34
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  233 South Wacker Drive/6300 Sears Tower
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/039,198A
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rin-Laures, Li-Hsien
;      REGISTRATION NUMBER:  33,547
;      REFERENCE/DOCKET NUMBER:  27856/34391
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 474-6300
;      TELEFAX:  (312) 474-0448
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  373 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-039-198A-14

  Query Match            93.2%;  Score 41;  DB 3;  Length 373;
  Best Local Similarity  88.9%;  Pred. No. 0.88;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FGRSFTLAS 9
              :|||||||||
Db        246 YGRSFTLAS 254


RESULT 10
US-09-039-198A-15
; Sequence 15, Application US/09039198A
; Patent No. 6200951
;  GENERAL INFORMATION:
;    APPLICANT:  Gray, Patrick W.
;    APPLICANT:  Tjoelker, Larry W.
;    TITLE OF INVENTION:  CHITINASE CHITIN-BINDING FRAGMENTS
;    NUMBER OF SEQUENCES:  34
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  233 South Wacker Drive/6300 Sears Tower
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/039,198A
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rin-Laures, Li-Hsien
;      REGISTRATION NUMBER:  33,547
;      REFERENCE/DOCKET NUMBER:  27866/34391
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 474-6300
;      TELEFAX:  (312) 474-0448
;  INFORMATION FOR SEQ ID NO:  15:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  373 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-039-198A-15

  Query Match            93.2%;  Score 41;  DB 3;  Length 373;
  Best Local Similarity  88.9%;  Pred. No. 0.88;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FGRSFTLAS 9
              :|||||||||
Db        246 YGRSFTLAS 254


RESULT 11
US-08-877-599-14
; Sequence 14, Application US/08877599
; Patent No. 6372212
;  GENERAL INFORMATION:
;    APPLICANT:  Gray, Patrick W.
;    TITLE OF INVENTION:  Chitinase Materials and Methods
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
```

```
;      APPLICATION NUMBER:  US/08/877,599
;      FILING DATE:
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/663,618
;      FILING DATE:  14-JUN-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rin-Laures, Li-Hsien
;      REGISTRATION NUMBER:  33,547
;      REFERENCE/DOCKET NUMBER:  27866/33994
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  373 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-877-599-14

  Query Match             93.2%;  Score 41;  DB 4;  Length 373;
  Best Local Similarity   88.9%;  Pred. No. 0.88;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FGRSFTLAS 9
              :|||||||||
Db        246 YGRSFTLAS 254


RESULT 12
US-08-877-599-15
; Sequence 15, Application US/08877599
; Patent No. 6372212
;  GENERAL INFORMATION:
;    APPLICANT:  Gray, Patrick W.
;    TITLE OF INVENTION:  Chitinase Materials and Methods
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/877,599
;      FILING DATE:
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/663,618
;      FILING DATE:  14-JUN-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rin-Laures, Li-Hsien
;      REGISTRATION NUMBER:  33,547
;      REFERENCE/DOCKET NUMBER:  27866/33994
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  15:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  373 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-877-599-15

  Query Match             93.2%;  Score 41;  DB 4;  Length 373;
  Best Local Similarity   88.9%;  Pred. No. 0.88;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FGRSFTLAS 9
              :|||||||||
Db        246 YGRSFTLAS 254


RESULT 13
US-09-267-574-14
; Sequence 14, Application US/09267574
; Patent No. 6399571
; GENERAL INFORMATION:
;  APPLICANT: Gray, Patrick W.
;  APPLICANT: Tjoelker, Larry W.
;  TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
;  FILE REFERENCE: 27866/35407
;  CURRENT APPLICATION NUMBER: US/09/267,574
;  CURRENT FILING DATE: 1999-03-12
;  EARLIER APPLICATION NUMBER: 09/039,198
;  EARLIER FILING DATE: 1998-03-12
;  NUMBER OF SEQ ID NOS: 39
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
;   LENGTH: 373
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-267-574-14

  Query Match             93.2%;  Score 41;  DB 4;  Length 373;
  Best Local Similarity   88.9%;  Pred. No. 0.88;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FGRSFTLAS 9
              :|||||||||
Db        246 YGRSFTLAS 254


RESULT 14
US-09-267-574-15
; Sequence 15, Application US/09267574
; Patent No. 6399571
; GENERAL INFORMATION:
;  APPLICANT: Gray, Patrick W.
;  APPLICANT: Tjoelker, Larry W.
;  TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
;  FILE REFERENCE: 27866/35407
;  CURRENT APPLICATION NUMBER: US/09/267,574
;  CURRENT FILING DATE: 1999-03-12
;  EARLIER APPLICATION NUMBER: 09/039,198
;  EARLIER FILING DATE: 1998-03-12
;  NUMBER OF SEQ ID NOS: 39
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
;   LENGTH: 373
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-267-574-15

  Query Match             93.2%;  Score 41;  DB 4;  Length 373;
  Best Local Similarity   88.9%;  Pred. No. 0.88;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FGRSFTLAS 9
              :|||||||||
Db        246 YGRSFTLAS 254


RESULT 15
```

```
US-08-486-839-6
; Sequence 6, Application US/08486839
; Patent No. 5928928
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  A human chitinase, its recombinant
;    TITLE OF INVENTION:  production, its use for decomposing chitin, its use
;    TITLE OF INVENTION:  in therapy or prophylaxis against infection diseases.
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hoffmann & Baron
;      STREET:  350 Jericho Turnpike
;      CITY:  Jericho
;      STATE:  New York
;      COUNTRY:  United States of America
;      ZIP:  11758
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;      SOFTWARE:
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/486,839
;      FILING DATE:  07 - June - 1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Baron, Ronald J.
;      REGISTRATION NUMBER:  29,281
;      REFERENCE/DOCKET NUMBER:  294-26
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (516) 822-3550
;      TELEFAX:  (516) 822-3582
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  387 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
US-08-486-839-6

  Query Match            93.2%;  Score 41;  DB 2;  Length 387;
  Best Local Similarity   88.9%;  Pred. No. 0.92;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PGRSFTLAS 9
              :|.||||:
Db        267 YGRSFTLAS 275


Search completed: September 24, 2003, 13:51:12
Job time : 11.9213 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 24, 2003, 13:47:40 ; Search time 20.1236 Seconds
(without alignments)
67.670 Million cell updates/sec

Title: US-09-744-282-1
Perfect score: 44
Sequence: 1 FGRSFILAS 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 566894 seqs, 151307093 residues

Total number of hits satisfying chosen parameters: 566894

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	44	100.0	13	10	US-09-459-749D-5	Sequence 5, Appli
2	44	100.0	15	10	US-09-985-065-2	Sequence 2, Appli
3	44	100.0	383	10	US-09-459-749D-17	Sequence 17, Appl
4	44	100.0	383	15	US-10-097-340-45	Sequence 45, Appl
5	41	93.2	373	12	US-10-161-547-14	Sequence 14, Appl
6	41	93.2	373	12	US-10-161-547-15	Sequence 15, Appl
7	41	93.2	445	15	US-10-004-219B-10	Sequence 10, Appl
8	41	93.2	466	12	US-10-161-547-2	Sequence 2, Appli
9	41	93.2	466	12	US-10-161-547-4	Sequence 4, Appli
10	38	86.4	16	10	US-09-985-065-3	Sequence 3, Appli
11	38	86.4	199	9	US-09-815-837-56	Sequence 56, Appl
12	38	86.4	199	9	US-09-815-837-60	Sequence 60, Appl
13	38	86.4	208	9	US-09-815-837-102	Sequence 102, App
14	38	86.4	210	9	US-09-815-837-40	Sequence 40, Appl
15	38	86.4	211	9	US-09-815-837-39	Sequence 39, Appl
16	38	86.4	211	9	US-09-815-837-59	Sequence 59, Appl
17	38	86.4	212	9	US-09-815-837-37	Sequence 37, Appl
18	38	86.4	212	9	US-09-815-837-38	Sequence 38, Appl
19	38	86.4	213	9	US-09-815-837-41	Sequence 41, Appl
20	38	86.4	213	9	US-09-815-837-43	Sequence 43, Appl
21	38	86.4	213	9	US-09-815-837-44	Sequence 44, Appl
22	38	86.4	213	9	US-09-815-837-57	Sequence 57, Appl
23	38	86.4	214	9	US-09-815-837-42	Sequence 42, Appl
24	38	86.4	215	9	US-09-815-837-45	Sequence 45, Appl
25	38	86.4	216	9	US-09-815-837-46	Sequence 46, Appl
26	38	86.4	220	9	US-09-815-837-47	Sequence 47, Appl
27	38	86.4	220	9	US-09-815-837-48	Sequence 48, Appl
28	32	72.7	13	10	US-09-985-065-1	Sequence 1, Appli
29	32	72.7	190	10	US-09-893-737-198	Sequence 198, App
30	32	72.7	209	9	US-09-815-837-58	Sequence 58, Appl
31	32	72.7	554	15	US-10-184-128-2	Sequence 2, Appli
32	31	70.5	168	15	US-10-219-220-69	Sequence 69, Appl
33	31	70.5	286	12	US-10-239-663-48	Sequence 48, Appl
34	30	68.2	36	9	US-09-864-761-48719	Sequence 48719, A
35	30	68.2	104	9	US-09-925-301-1602	Sequence 1602, Ap
36	30	68.2	524	10	US-09-738-626-3502	Sequence 3502, Ap
37	30	68.2	1118	12	US-10-154-419-58	Sequence 58, Appl
38	30	68.2	1118	12	US-10-146-733-53	Sequence 53, Appl
39	30	68.2	1118	14	US-10-024-623-8	Sequence 8, Appli
40	29	65.9	88	15	US-10-196-698-4803	Sequence 4803, Ap
41	29	65.9	132	10	US-09-764-864-1333	Sequence 1333, Ap
42	29	65.9	135	10	US-09-764-864-1098	Sequence 1098, Ap
43	29	65.9	142	10	US-09-764-864-1518	Sequence 1518, Ap
44	29	65.9	199	15	US-10-156-781-8757	Sequence 8757, Ap
45	29	65.9	200	15	US-10-211-962-107	Sequence 107, App

ALIGNMENTS

RESULT 1
US-09-459-749D-5
; Sequence 5, Application US/09459749D
; Patent No. US20020136716A1
; GENERAL INFORMATION:
; APPLICANT: Millis, Albert J. T.
; TITLE OF INVENTION: Compositions and Methods For Altering Cell Migration
; FILE REFERENCE: 0794.016A
; CURRENT APPLICATION NUMBER: US/09/459,749D
; CURRENT FILING DATE: 1999-12-10
; PRIOR APPLICATION NUMBER: 60/111,856
; PRIOR FILING DATE: 1998-12-11
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:SITE
; NAME/KEY: SITE
; LOCATION: (1)..(13)
; OTHER INFORMATION: Peptide corresponding to amino acids 261- 273 of
; OTHER INFORMATION: gp38k protein
US-09-459-749D-5

Query Match 100.0%; Score 44; DB 10; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.018;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FGRSFILAS 9
||||||||| |
Db 1 FGRSFILAS 9

RESULT 2
US-09-985-065-2
; Sequence 2, Application US/09985065

```
; Patent No. US20020143150A1
;  GENERAL INFORMATION:
;       APPLICANT: Stoenbakkers PGA
;       TITLE OF INVENTION: Method of preparing a monoclonal
;                           antibody (as amended)
;       NUMBER OF SEQUENCES: 3
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Akzo No. US20020143150A1el Patent Dept.
;            STREET: 1300 Piccard Drive, Suite 206
;            CITY: Rockville
;            STATE: Maryland
;            COUNTRY: US
;            ZIP: 20850
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/985,065
;            FILING DATE: 01-No. US20020143150A1-2001
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 09/405,745
;            FILING DATE: 01-NOV-2001
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Sullivan, Michael G.
;            REGISTRATION NUMBER: 35,377
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 301-948-7400
;            TELEFAX: 301-948-9751
;   INFORMATION FOR SEQ ID NO: 2:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 15 amino acids
;            TYPE: amino acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       FRAGMENT TYPE: internal
;       SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-985-065-2

  Query Match           100.0%;  Score 44;  DB 10;  Length 15;
  Best Local Similarity  100.0%;  Pred. No. 0.02;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FGRSFTLAS 9
              |||||||||
Db          1 FGRSFTLAS 9


RESULT 3
US-09-459-749D-17
; Sequence 17, Application US/09459749D
; Patent No. US20020136716A1
; GENERAL INFORMATION:
;  APPLICANT: Millis, Albert J. T.
;  TITLE OF INVENTION: Compositions and Methods For Altering Cell Migration
;  FILE REFERENCE: 0794.016A
;  CURRENT APPLICATION NUMBER: US/09/459,749D
;  CURRENT FILING DATE:  1999-12-10
;  PRIOR APPLICATION NUMBER: 60/111,856
;  PRIOR FILING DATE: 1998-12-11
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
;   LENGTH: 383
;   TYPE: PRT
;   ORGANISM: Sus scrofa
US-09-459-749D-17

  Query Match           100.0%;  Score 44;  DB 10;  Length 383;
  Best Local Similarity  100.0%;  Pred. No. 0.62;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FGRSFTLAS 9
              |||||||||
Db        261 FGRSFTLAS 269


RESULT 4
US-10-097-340-45
; Sequence 45, Application US/10097340
; Publication No. US20030087250A1
; GENERAL INFORMATION:
;  APPLICANT: John MONAHAN
;  APPLICANT:  Manjula GANNAVARAPU
;  APPLICANT:  Sebastian HOERSCH
;  APPLICANT:  Shubhangi KAMATKAR
;  APPLICANT:  Steve G. KOVATS
;  APPLICANT:  Rachel E. MEYERS
;  APPLICANT:  Michael MORRISEY
;  APPLICANT:  Peter OLANDT
;  APPLICANT:  Ami SEN
;  APPLICANT:  Peter VEIBY
;  APPLICANT:  Gordon B. MILLS
;  APPLICANT:  Robert C. BAST, Jr.
;  APPLICANT:  Karen LU
;  APPLICANT:  Rosemarie SCHMANDT
;  APPLICANT:  Xumei ZHAO
;  APPLICANT:  Karen GLATT
;  TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
;  TITLE OF INVENTION:  Assessment, Prevention, and Therapy of Ovarian Cancer
;  FILE REFERENCE: MRI-030
;  CURRENT APPLICATION NUMBER: US/10/097,340
;  CURRENT FILING DATE:  2002-03-14
;  PRIOR APPLICATION NUMBER: 60/276,025
;  PRIOR FILING DATE: 2001-03-14
;  PRIOR APPLICATION NUMBER: 60/325,149
;  PRIOR FILING DATE: 2001-09-26
;  PRIOR APPLICATION NUMBER: 60/276,026
;  PRIOR FILING DATE: 2001-03-14
;  PRIOR APPLICATION NUMBER: 60/324,967
;  PRIOR FILING DATE: 2001/09/26
;  PRIOR APPLICATION NUMBER: 60/311,732
;  PRIOR FILING DATE: 2001-08-10
;  PRIOR APPLICATION NUMBER: 60/325,102
;  PRIOR FILING DATE: 2001-09-26
;  PRIOR APPLICATION NUMBER: 60/323,580
;  PRIOR FILING DATE: 2001-09-19
;  NUMBER OF SEQ ID NOS: 363
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 45
;   LENGTH: 383
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-097-340-45

  Query Match           100.0%;  Score 44;  DB 15;  Length 383;
  Best Local Similarity  100.0%;  Pred. No. 0.62;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FGRSFTLAS 9
              |||||||||
Db        261 FGRSFTLAS 269


RESULT 5
US-10-161-547-14
; Sequence 14, Application US/10161547
; Publication No. US20030143216A1
; GENERAL INFORMATION:
;  APPLICANT: Gray, Patrick W.
;  APPLICANT:  Tjoelker, Larry W.
;  TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
;  FILE REFERENCE: 27866/35407
```

```
;  CURRENT APPLICATION NUMBER: US/10/161,547
;  CURRENT FILING DATE:  2002-06-03
;  PRIOR APPLICATION NUMBER: US/09/267,574
;  PRIOR FILING DATE: 1999-03-12
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/039,198
;  PRIOR FILING DATE: EARLIER FILING DATE: 1998-03-12
;  NUMBER OF SEQ ID NOS: 39
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
;   LENGTH: 373
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-161-547-14

  Query Match           93.2%;  Score 41;  DB 12;  Length 373;
  Best Local Similarity  88.9%;  Pred. No. 2.4;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FGRSFTLAS 9
              :|||||||:|
Db        246 YGRSFTLAS 254


RESULT 6
US-10-161-547-15
; Sequence 15, Application US/10161547
; Publication No. US20030143216A1
; GENERAL INFORMATION:
;  APPLICANT: Gray, Patrick W.
;  APPLICANT:  Tjoelker, Larry W.
;  TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
;  FILE REFERENCE: 27866/35407
;  CURRENT APPLICATION NUMBER: US/10/161,547
;  CURRENT FILING DATE:  2002-06-03
;  PRIOR APPLICATION NUMBER: US/09/267,574
;  PRIOR FILING DATE: 1999-03-12
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/039,198
;  PRIOR FILING DATE: EARLIER FILING DATE: 1998-03-12
;  NUMBER OF SEQ ID NOS: 39
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
;   LENGTH: 373
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-161-547-15

  Query Match           93.2%;  Score 41;  DB 12;  Length 373;
  Best Local Similarity  88.9%;  Pred. No. 2.4;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FGRSFTLAS 9
              :|||||||:|
Db        246 YGRSFTLAS 254


RESULT 7
US-10-004-219B-10
; Sequence 10, Application US/10004219B
; Publication No. US20030087414A1
; GENERAL INFORMATION:
;  APPLICANT: Macrozyme
;  APPLICANT:  Aerts, Johannes M.F.G.
;  APPLICANT:  Boot, Rolf G.
;  TITLE OF INVENTION: A mammalian mucinase, its recombinant production, and
;  TITLE OF INVENTION:  its use in therapy or prophylaxis against diseases in
;  TITLE OF INVENTION:  which mucus is involved or infection diseases
;  FILE REFERENCE: 2183-5136US
;  CURRENT APPLICATION NUMBER: US/10/004,219B
;  CURRENT FILING DATE:  2001-11-02
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
;   LENGTH: 445
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (1)..(445)
;   OTHER INFORMATION: /note="Human chitotriosidase"
US-10-004-219B-10

  Query Match           93.2%;  Score 41;  DB 15;  Length 445;
  Best Local Similarity  88.9%;  Pred. No. 2.9;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FGRSFTLAS 9
              :|||||||:|
Db        246 YGRSFTLAS 254


RESULT 8
US-10-161-547-2
; Sequence 2, Application US/10161547
; Publication No. US20030143216A1
; GENERAL INFORMATION:
;  APPLICANT: Gray, Patrick W.
;  APPLICANT:  Tjoelker, Larry W.
;  TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
;  FILE REFERENCE: 27866/35407
;  CURRENT APPLICATION NUMBER: US/10/161,547
;  CURRENT FILING DATE:  2002-06-03
;  PRIOR APPLICATION NUMBER: US/09/267,574
;  PRIOR FILING DATE: 1999-03-12
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/039,198
;  PRIOR FILING DATE: EARLIER FILING DATE: 1998-03-12
;  NUMBER OF SEQ ID NOS: 39
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 466
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-161-547-2

  Query Match           93.2%;  Score 41;  DB 12;  Length 466;
  Best Local Similarity  88.9%;  Pred. No. 3;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FGRSFTLAS 9
              :|||||||:|
Db        267 YGRSFTLAS 275


RESULT 9
US-10-161-547-4
; Sequence 4, Application US/10161547
; Publication No. US20030143216A1
; GENERAL INFORMATION:
;  APPLICANT: Gray, Patrick W.
;  APPLICANT:  Tjoelker, Larry W.
;  TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
;  FILE REFERENCE: 27866/35407
;  CURRENT APPLICATION NUMBER: US/10/161,547
;  CURRENT FILING DATE:  2002-06-03
;  PRIOR APPLICATION NUMBER: US/09/267,574
;  PRIOR FILING DATE: 1999-03-12
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/039,198
;  PRIOR FILING DATE: EARLIER FILING DATE: 1998-03-12
;  NUMBER OF SEQ ID NOS: 39
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
;   LENGTH: 466
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-161-547-4

  Query Match           93.2%;  Score 41;  DB 12;  Length 466;
```

```
  Best Local Similarity   88.9%;  Pred. No. 3;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FGRSFTLAS 9
              :|:||||||
Db        267 YGRSFTLAS 275


RESULT 10
US-09-985-065-3
; Sequence 3, Application US/09985065
; Patent No. US20020143150A1
;   GENERAL INFORMATION:
;        APPLICANT: Steenbakkers PGA
;        TITLE OF INVENTION: Method of preparing a monoclonal
;                            antibody (as amended)
;        NUMBER OF SEQUENCES: 3
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Akzo No. US20020143150A1o1 Patent Dept.
;             STREET: 1300 Piccard Drive, Suite 206
;             CITY: Rockville
;             STATE: Maryland
;             COUNTRY: US
;             ZIP: 20850
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/985.065
;             FILING DATE: 01-No. US20020143150A1-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/405,745
;             FILING DATE: 01-NOV-2001
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Sullivan, Michael G.
;             REGISTRATION NUMBER: 35,377
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 301-948-7400
;             TELEFAX: 301-948-9751
;   INFORMATION FOR SEQ ID NO: 3:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 16 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        FRAGMENT TYPE: internal
;        SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-985-065-3

  Query Match             86.4%;  Score 38;  DB 10;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 0.35;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 GRSFTLAS 9
              ||||||||
Db          1 GRSFTLAS 8


RESULT 11
US-09-815-837-56
; Sequence 56, Application US/09815837
; Patent No. US20020082411A1
; GENERAL INFORMATION:
;  APPLICANT: Carter, Darrick
;  APPLICANT:  Zhu, Shirley
;  APPLICANT:  Arimilli, Subhashini
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Immune Medators and Related Methods
;  FILE REFERENCE: 014058-005670US
;  CURRENT APPLICATION NUMBER: US/09/815,837
;  CURRENT FILING DATE:  2001-03-22
;  PRIOR APPLICATION NUMBER: US 60/191,274
;  PRIOR FILING DATE: 2000-03-22
;  PRIOR APPLICATION NUMBER: US 60/204,249
;  PRIOR FILING DATE: 2000-05-15
;  PRIOR APPLICATION NUMBER: US 60/264,003
;  PRIOR FILING DATE: 2001-01-23
;  NUMBER OF SEQ ID NOS: 129
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 56
;   LENGTH: 199
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:CO528-AC single
;   OTHER INFORMATION: chain recombinant MHC class II molecule
US-09-815-837-56

  Query Match             86.4%;  Score 38;  DB 9;  Length 199;
  Best Local Similarity   100.0%;  Pred. No. 5;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 GRSFTLAS 9
              ||||||||
Db          5 GRSFTLAS 12


RESULT 12
US-09-815-837-60
; Sequence 60, Application US/09815837
; Patent No. US20020082411A1
; GENERAL INFORMATION:
;  APPLICANT: Carter, Darrick
;  APPLICANT:  Zhu, Shirley
;  APPLICANT:  Arimilli, Subhashini
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Immune Medators and Related Methods
;  FILE REFERENCE: 014058-005670US
;  CURRENT APPLICATION NUMBER: US/09/815,837
;  CURRENT FILING DATE:  2001-03-22
;  PRIOR APPLICATION NUMBER: US 60/191,274
;  PRIOR FILING DATE: 2000-03-22
;  PRIOR APPLICATION NUMBER: US 60/204,249
;  PRIOR FILING DATE: 2000-05-15
;  PRIOR APPLICATION NUMBER: US 60/264,003
;  PRIOR FILING DATE: 2001-01-23
;  NUMBER OF SEQ ID NOS: 129
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 60
;   LENGTH: 199
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:CO528 single
;   OTHER INFORMATION: chain recombinant MHC class II molecule
US-09-815-837-60

  Query Match             86.4%;  Score 38;  DB 9;  Length 199;
  Best Local Similarity   100.0%;  Pred. No. 5;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 GRSFTLAS 9
              ||||||||
Db          5 GRSFTLAS 12


RESULT 13
US-09-815-837-102
; Sequence 102, Application US/09815837
; Patent No. US20020082411A1
```

```
; GENERAL INFORMATION:
;  APPLICANT: Carter, Darrick
;  APPLICANT:  Zhu, Shirley
;  APPLICANT:  Arimilli, Subhashini
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Immune Medators and Related Methods
;  FILE REFERENCE: 014058-005670US
;  CURRENT APPLICATION NUMBER: US/09/815,837
;  CURRENT FILING DATE:  2001-03-22
;  PRIOR APPLICATION NUMBER: US 60/191,274
;  PRIOR FILING DATE: 2000-03-22
;  PRIOR APPLICATION NUMBER: US 60/204,249
;  PRIOR FILING DATE: 2000-05-15
;  PRIOR APPLICATION NUMBER: US 60/264,003
;  PRIOR FILING DATE: 2001-01-23
;  NUMBER OF SEQ ID NOS: 129
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 102
;   LENGTH: 208
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial
;   OTHER INFORMATION: Sequence:gp39-beta1/alpha1 HLA-DR4 CO563 human
;   OTHER INFORMATION: single chain molecule
US-09-815-837-102

  Query Match            86.4%;  Score 38;  DB 9;  Length 208;
  Best Local Similarity  100.0%;  Pred. No. 5.2;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 GRSFTLAS 9
              ||||||||
Db          5 GRSFTLAS 12


RESULT 14
US-09-815-837-40
; Sequence 40, Application US/09815837
; Patent No. US20020082411A1
; GENERAL INFORMATION:
;  APPLICANT: Carter, Darrick
;  APPLICANT:  Zhu, Shirley
;  APPLICANT:  Arimilli, Subhashini
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Immune Medators and Related Methods
;  FILE REFERENCE: 014058-005670US
;  CURRENT APPLICATION NUMBER: US/09/815,837
;  CURRENT FILING DATE:  2001-03-22
;  PRIOR APPLICATION NUMBER: US 60/191,274
;  PRIOR FILING DATE: 2000-03-22
;  PRIOR APPLICATION NUMBER: US 60/204,249
;  PRIOR FILING DATE: 2000-05-15
;  PRIOR APPLICATION NUMBER: US 60/264,003
;  PRIOR FILING DATE: 2001-01-23
;  NUMBER OF SEQ ID NOS: 129
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 40
;   LENGTH: 210
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:CO595 single
;   OTHER INFORMATION: chain recombinant MHC class II molecule
US-09-815-837-40

  Query Match            86.4%;  Score 38;  DB 9;  Length 210;
  Best Local Similarity  100.0%;  Pred. No. 5.3;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 GRSFTLAS 9
              ||||||||
Db          5 GRSFTLAS 12


RESULT 15
US-09-815-837-39
; Sequence 39, Application US/09815837
; Patent No. US20020082411A1
; GENERAL INFORMATION:
;  APPLICANT: Carter, Darrick
;  APPLICANT:  Zhu, Shirley
;  APPLICANT:  Arimilli, Subhashini
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Immune Medators and Related Methods
;  FILE REFERENCE: 014058-005670US
;  CURRENT APPLICATION NUMBER: US/09/815,837
;  CURRENT FILING DATE:  2001-03-22
;  PRIOR APPLICATION NUMBER: US 60/191,274
;  PRIOR FILING DATE: 2000-03-22
;  PRIOR APPLICATION NUMBER: US 60/204,249
;  PRIOR FILING DATE: 2000-05-15
;  PRIOR APPLICATION NUMBER: US 60/264,003
;  PRIOR FILING DATE: 2001-01-23
;  NUMBER OF SEQ ID NOS: 129
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 39
;   LENGTH: 211
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:CO596 single
;   OTHER INFORMATION: chain recombinant MHC class II molecule
US-09-815-837-39

  Query Match            86.4%;  Score 38;  DB 9;  Length 211;
  Best Local Similarity  100.0%;  Pred. No. 5.3;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 GRSFTLAS 9
              ||||||||
Db          5 GRSFTLAS 12


Search completed: September 24, 2003, 14:22:25
Job time : 21.1236 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 24, 2003, 13:34:08 ; Search time 10.9213 Seconds
(without alignments)
79.250 Million cell updates/sec

Title: US-09-744-282-1
Perfect score: 44
Sequence: 1 FGRSFTLAS 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	44	100.0	383	2	S51327	heparin-binding gl
2	44	100.0	383	2	A49562	cartilage glycopro
3	41	93.2	405	2	S61551	breast-regressing
4	38	86.4	525	2	T44445	chitinase (EC 3.2.
5	36	81.8	203	2	C25511	Cc protein - fruit
6	36	81.8	1074	2	F72217	conserved hypothet
7	35	79.5	554	2	A56598	chitinase (EC 3.2.
8	35	79.5	765	2	E84553	hypothetical prote
9	34	77.3	314	2	B83929	response regulator
10	33	75.0	267	2	F82701	conserved hypothet
11	33	75.0	684	2	B69308	conserved hypothet
12	33	75.0	1635	2	T14075	chitinase (EC 3.2.
13	32	72.7	114	2	D71048	hypothetical prote
14	32	72.7	307	2	AH3112	hypothetical prote
15	32	72.7	307	2	E98174	oligopeptide ABC t
16	32	72.7	333	2	T50753	protochlorophyllid
17	32	72.7	333	2	S17823	protochlorophyllid
18	32	72.7	333	2	S30915	protochlorophyllid
19	32	72.7	334	2	T50882	chlorophillide red
20	32	72.7	405	2	T16922	hypothetical prote
21	32	72.7	415	2	S34140	DNA-binding protei
22	32	72.7	415	2	S37690	DNA-binding protei
23	32	72.7	483	2	A53918	chitinase (EC 3.2.
24	32	72.7	592	2	T42078	serine/threonine p
25	32	72.7	731	2	T34217	hypothetical prote
26	32	72.7	1033	2	T37715	actin-interacting
27	31	70.5	246	2	T20341	hypothetical prote
28	31	70.5	515	2	C95899	probable didpeptid
29	31	70.5	325	2	138616	zinc finger protei
30	31	70.5	565	2	T38941	RNA binding protei
31	31	70.5	573	2	JQ0135	hypothetical 62.8K
32	31	70.5	716	2	AC2449	ABC transporter AT
33	31	70.5	1733	2	S27939	tensin - chicken
34	31	70.5	1744	2	A54970	tensin, cardiac mu
35	31	70.5	1792	2	A57075	tensin - chicken (
36	30	68.2	114	2	F87148	hypothetical prote
37	30	68.2	179	2	AG1994	hypothetical prote
38	30	68.2	196	2	H72551	hypothetical prote
39	30	68.2	199	1	S50398	ribosomal protein
40	30	68.2	199	2	S67618	ribosomal protein
41	30	68.2	202	2	T52146	ribosomal protein
42	30	68.2	292	2	S77168	hypothetical prote
43	30	68.2	338	2	AE3079	hypothetical prote
44	30	68.2	376	2	G95949	probable oligopept
45	30	68.2	381	2	AB0734	probable bacteriop

ALIGNMENTS

RESULT 1
S51327
heparin-binding glycoprotein 38K - pig
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 22-Jun-1999
C;Accession: S51327
R;Shackelton, L.M.; Mann, D.M.; Millis, A.J.T.
submitted to the EMBL Data Library, January 1995
A;Description: Identification of a 38kDa heparin-binding glycoprotein (gp38k) in dif
A;Reference number: S51327
A;Accession: S51327
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-383 <SHA>
A;Cross-references: EMBL:Z47803; NID:g634097; PIDN:CAA87764.1; PID:g634098
C;Superfamily: Streptomyces chitinase chi40

Query Match 100.0%; Score 44; DB 2; Length 383;
Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FGRSFTLAS 9
|||||||||
Db 261 FGRSFTLAS 269

RESULT 2
A49562
cartilage glycoprotein gp39 precursor - human
N;Alternate names: 39K synovial protein
C;Species: Homo sapiens (man)
C;Date: 23-Mar-1995 #sequence_revision 23-Mar-1995 #text_change 22-Jun-1999
C;Accession: A49562; S10677; A33162
R;Hakala, B.E.; White, C.; Recklies, A.D.
J. Biol. Chem. 268, 25803-25810, 1993
A;Title: Human cartilage gp-39, a major secretory product of articular chondrocytes
A;Reference number: A49562; MUID:94064658; PMID:8245017
A;Accession: A49562
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-383 <HAK>
A;Cross-references: GB:M80927; NID:g348911; PIDN:AAA16074.1; PID:g348912
R;Nyirkos, P.; Golds, E.E.
Biochem. J. 269, 265-268, 1990
A;Title: Human synovial cells secrete a 39 kDa protein similar to a bovine mammary p
A;Reference number: S10677; MUID:90328983; PMID:2375755
A;Accession: S10677
A;Molecule type: protein
A;Residues: 22-40,'X',42-45 <NY2>
C;Superfamily: Streptomyces chitinase chi40
C;Keywords: cartilage; extracellular protein; glycoprotein
F;1-21/Domain: signal sequence #status predicted <SIG>

```
F;22-383/Product: cartilage glycoprotein gp39 #status predicted <MAT>

  Query Match           100.0%;  Score 44;  DB 2;  Length 383;
  Best Local Similarity  100.0%;  Pred. No. 0.14;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FGRSFTLAS 9
              |||||||||
Db        261 FGRSFTLAS 269


RESULT 3
S61551
breast-regressing protein brp39 precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
C;Accession: S61551; S61550; I48271
R;Morrison, B.W.; Leder, P.
Oncogene 9, 3417-3426, 1994
A;Title: neu and ras initiate murine mammary tumors that share genetic markers generally
A;Reference number: I48271; MUID:95060797; PMID:7970700
A;Accession: S61551
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-405 <MOR1>
A;Cross-references: EMBL:X93035; NID:g1085065; PIDN:CAA63603.1; PID:g1085069
R;Morrison, B.W.
submitted to the EMBL Data Library, November 1995
A;Reference number: S61550
A;Accession: S61550
A;Molecule type: mRNA
A;Residues: 1-245,'I',247-330,'H',332-350,'MVWALDLDDFQGTCQPKEFFPLTNAIKDALA' <MOR2>
A;Cross-references: EMBL:X93035; NID:g1085065; PIDN:CAA63603.1; PID:g1085066
A;Note: the differences at the carboxyl end are due to a frameshift error
C;Genetics:
A;Gene: brp39
C;Superfamily: Streptomyces chitinase chi40
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-405/Product: breast-regressing protein brp39 #status predicted <MAT>

  Query Match           93.2%;  Score 41;  DB 2;  Length 405;
  Best Local Similarity  88.9%;  Pred. No. 0.65;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FGRSFTLAS 9
              |:|||||:|
Db        262 FGKSFTLAS 270


RESULT 4
T44445
chitinase (EC 3.2.1.14) [imported] - African malaria mosquito
C;Species: Anopheles gambiae (African malaria mosquito)
C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jan-2000
C;Accession: T44445
R;Shen, Z.; Jacobs-Lorena, M.
submitted to the EMBL Data Library, June 1997
A;Reference number: Z22771
A;Accession: T44445
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-525 <SHE>
A;Cross-references: EMBL:AF008575; PIDN:AAB87754.1
A;Experimental source: adult; gut
C;Genetics:
A;Gene: chi-1
C;Keywords: glycosidase; hydrolase

  Query Match           86.4%;  Score 38;  DB 2;  Length 525;
  Best Local Similarity  77.8%;  Pred. No. 3.7;
  Matches    7;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FGRSFTLAS 9
              :|:|:||||
Db        279 YGRNFTLAS 287


RESULT 5
C25511
Cc protein - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 21-Jul-2000
C;Accession: C25511
R;Eveleth Jr., D.D.; Marsh, J.L.
Nucleic Acids Res. 14, 6169-6183, 1986
A;Title: Sequence and expression of the Cc gene, a member of the dopa decarboxylase g
A;Reference number: A93634; MUID:86312687; PMID:3092183
A;Accession: C25511
A;Molecule type: DNA
A;Residues: 1-203 <EVE>
A;Cross-references: GB:X04227; NID:g7695; PIDN:CAA27807.1; PID:g295740
C;Genetics:
A;Gene: FlyBase:l(2)37Cc
A;Cross-references: FlyBase:FBgn0002031
A;Introns: 67/3

  Query Match           81.8%;  Score 36;  DB 2;  Length 203;
  Best Local Similarity  87.5%;  Pred. No. 3.9;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FGRSFTLA 8
              ||| ||||
Db         98 FGREFTLA 105


RESULT 6
F72217
conserved hypothetical protein - Thermotoga maritima (strain MSB8)
C;Species: Thermotoga maritima
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C;Accession: F72217
R;Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hic
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson,
 C.M.
Nature 399, 323-329, 1999
A;Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome
A;Reference number: A72200; MUID:99287316; PMID:10360571
A;Accession: F72217
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1074 <ARN>
A;Cross-references: GB:AE001812; GB:AE000512; NID:g4982302; PIDN:AAD36800.1; PID:g498
A;Experimental source: strain MSB8
C;Genetics:
A;Gene: TM1735

  Query Match           81.8%;  Score 36;  DB 2;  Length 1074;
  Best Local Similarity  77.8%;  Pred. No. 21;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FGRSFTLAS 9
              |||:|| |
Db       1038 FGRNFTLES 1046


RESULT 7
A56596
chitinase (EC 3.2.1.14) - tobacco hornworm
C;Species: Manduca sexta (tobacco hornworm)
C;Date: 11-Aug-1995 #sequence_revision 11-Aug-1995 #text_change 29-Jan-1999
C;Accession: A56596
R;Kramer, K.J.; Corpuz, L.; Choi, H.K.; Muthukrishnan, S.
Insect Biochem. Mol. Biol. 23, 691-701, 1993
A;Title: Sequence of a cDNA and expression of the gene encoding epidermal and gut chi
A;Reference number: A56596; MUID:93357793; PMID:8353525
A;Accession: A56596
```

A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-554 <KRA>
A;Cross-references: GB:U02270; GB:S64757; NID:g406046; PID:g406049
A;Experimental source: larvae
A;Note: sequence extracted from NCBI backbone (NCBIN:136417, NCBIP:136418)
C;Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 79.5%; Score 35; DB 2; Length 554;
Best Local Similarity 66.7%; Pred. No. 17;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FGRSFTLAS 9
:|||:|||::
Db 270 YGRSFTLSA 278

RESULT 8
E84553
hypothetical protein At2g17550 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C;Accession: E84553
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: E84553
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-765 <STO>
A;Cross-references: GB:AE002093; NID:g4914376; PIDN:AAD32912.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g17550
A;Map position: 2

Query Match 79.5%; Score 35; DB 2; Length 765;
Best Local Similarity 77.8%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FGRSFTLAS 9
:|. |||.:
Db 211 FGRDFTLKS 219

RESULT 9
B83929
response regulatory protein (sensory transduction system) BH2234 [imported] - Bacillus H
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 14-Apr-2003
C;Accession: B83929
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: B83929
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-314 <STO>
A;Cross-references: GB:AP001514; GB:BA000004; NID:g10174613; PIDN:BAB05953.1; GSPDB:GN0
A;Experimental source: strain C-125
C;Genetics:
A;Gene: BH2234
C;Superfamily: response regulator with a receiver and a diguanylate cyclase (GGDEF) doma

Query Match 77.3%; Score 34; DB 2; Length 314;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GRSFTLA 8
|||||||
Db 181 GRSFTLA 187

RESULT 10
F82701
conserved hypothetical protein XF1272 [imported] - Xylella fastidiosa (strain 9a5c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C;Accession: F82701
R;Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Se
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: F82701
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-267 <SIM>
A;Cross-references: GB:AE003961; GB:AE003849; NID:g9106254; PIDN:AAF84081.1; GSPDB:G
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, F
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carre
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
Submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; F
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; L
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martin
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C
, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri,
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sat
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Sil
M.; Tsuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF1272

Query Match 75.0%; Score 33; DB 2; Length 267;
Best Local Similarity 85.7%; Pred. No. 23;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FGRSFTL 7
:|||||
Db 126 YGRSFTL 132

RESULT 11
B69308
conserved hypothetical protein AF0466 - Archaeoglobus fulgidus
C;Species: Archaeoglobus fulgidus
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
C;Accession: B69308
R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Do
.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness,
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A;Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes
Smith, H.O.; Woese, C.R.; Venter, J.C.
A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing ar
A;Reference number: A69250; MUID:98049343; PMID:9389475
A;Accession: B69308
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-684 <KLE>
A;Cross-references: GB:AE001072; GB:AE000782; NID:g2689395; PIDN:AAB90771.1; PID:g2

Query Match 75.0%; Score 33; DB 2; Length 684;
Best Local Similarity 77.8%; Pred. No. 58;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FGRSFTLAS 9
|| |||||
Db 396 FGLGFTLAS 404

```
RESULT 12
T14075
chitinase (EC 3.2.1.14) - yellow fever mosquito
C;Species: Aedes aegypti (yellow fever mosquito)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C;Accession: T14075
R;de la Vega, H.; Specht, C.A.; Liu, Y.; Robbins, P.W.
Insect Mol. Biol. 7, 233-239, 1997
A;Title: Chitinases are a multi-gene family in Aedes, Anopheles, and Drosophila.
A;Reference number: Z17872
A;Accession: T14075
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1635 <DEL>
A;Cross-references: EMBL:AF026492; NID:g2564720; PID:g2564721; PIDN:AAB81850.1
C;Genetics:
A;Gene: CHT2
A;Introns: 462/3; 524/3; 618/1; 951/3; 1151/2
C;Keywords: glycosidase; hydrolase; polysaccharide degradation

  Query Match             75.0%;  Score 33;  DB 2;  Length 1635;
  Best Local Similarity   75.0%;  Pred. No. 1.4e+02;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FGRSFTLA 8
              :|:|||:|
Db        486 YGQSFTLA 493


RESULT 13
D71048
hypothetical protein PH1676 - Pyrococcus horikoshii
C;Species: Pyrococcus horikoshii
C;Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 21-Jul-2000
C;Accession: D71048
R;Kawarabayasi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekin
 M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi
DNA Res. 5, 55-76, 1998
A;Title: Complete sequence and gene organization of the genome of a hyper-thermophilic a
A;Reference number: A71000; MUID:98344137; PMID:9679194
A;Accession: D71048
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-114 <KAW>
A;Cross-references: GB:AP000006; NID:g3236133; PIDN:BAA30788.1; PID:g3258105
A;Experimental source: strain OT3
A;Note: this accession replaces an interim accession for a sequence replaced by GenBank
C;Genetics:
A;Gene: PH1676

  Query Match             72.7%;  Score 32;  DB 2;  Length 114;
  Best Local Similarity   77.8%;  Pred. No. 16;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FGRSFTLAS 9
              | |||:|||
Db         12 FKRSFSLAS 20


RESULT 14
AH3112
hypothetical protein Atu4524 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
C;Accession: AH3112
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, J
erage, G.; Gillet, W.; Grant, C.; Guenthner, D.; Kutyavin, T.; Levy, R.; Li, M.; McClel
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; MUID:21608550; PMID:11743193
A;Accession: AH3112
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-307 <KUR>
A;Cross-references: GB:AE008689; PIDN:AAL45318.1; PID:g17743008; GSPDB:GN00187
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: Atu4524
A;Map position: linear chromosome

  Query Match             72.7%;  Score 32;  DB 2;  Length 307;
  Best Local Similarity   100.0%;  Pred. No. 43;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 FGRSFT 6
              ||||||
Db         78 FGRSFT 83


RESULT 15
F98174
oligopeptide ABC transporter (permease) (AP001508) [imported] - Agrobacterium tumefa
C;Species: Agrobacterium tumefaciens
C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 18-Nov-2002
C;Accession: F98174
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldm
 A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz,
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium
A;Reference number: A97359; MUID:21608551; PMID:11743194
A;Accession: F98174
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-307 <KUR>
A;Cross-references: GB:AE007870; PIDN:AAK88919.1; PID:g15158693; GSPDB:GN00170
C;Genetics:
A;Gene: AGR_L_689
A;Map position: linear chromosome

  Query Match             72.7%;  Score 32;  DB 2;  Length 307;
  Best Local Similarity   100.0%;  Pred. No. 43;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FGRSFT 6
              ||||||
Db         78 FGRSFT 83


Search completed: September 24, 2003, 13:49:17
Job time : 12.9213 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 24, 2003, 13:26:37 ; Search time 5.66292 Seconds
(without alignments)
74.739 Million cell updates/sec

Title: US-09-744-282-1
Perfect score: 44
Sequence: 1 PGRSFTLAS 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	44	100.0	383	1	C3L1_HUMAN	P36222 homo sapien
2	41	93.2	381	1	C3L1_MOUSE	Q61362 mus musculu
3	36	81.8	203	1	L2CC_DROME	P24156 drosophila
4	36	81.8	390	1	C3L2_HUMAN	Q15782 homo sapien
5	35	79.5	554	1	CH1T_MANSE	P36362 manduca sex
6	33	75.0	713	1	ZBT1_HUMAN	Q9y2k1 homo sapien
7	33	75.0	713	1	ZBT1_MOUSE	Q91v19 mus musculu
8	32	72.7	259	1	PYRK_METAC	P58886 methanosarc
9	32	72.7	333	1	BCHX_RHOCA	P26177 rhodobacter
10	32	72.7	333	1	BCHX_RHOSH	Q02431 rhodobacter
11	32	72.7	405	1	TRT_CAEEL	Q27371 caenorhabdi
12	32	72.7	415	1	ZN12_MICSA	P38621 micropterus
13	32	72.7	1385	1	FAT1_SCHPO	O13735 schizosacch
14	31	70.5	281	1	R40C_HUMAN	Q96s21 homo sapien
15	31	70.5	281	1	R40C_MOUSE	Q8vhq4 mus musculu
16	31	70.5	325	1	ZN36_HUMAN	P17029 homo sapien
17	31	70.5	1744	1	TENS_CHICK	Q04205 gallus gall
18	30	68.2	199	1	R13A_YEAST	Q12690 saccharomyc
19	30	68.2	199	1	R13B_YEAST	P40212 saccharomyc
20	30	68.2	202	1	RL13_CANAL	O59931 candida alb
21	30	68.2	446	1	ZN38_HUMAN	P17036 homo sapien
22	30	68.2	669	1	UVRB_SYNY3	Q55170 synechocyst
23	30	68.2	718	1	YB10_MYCPN	P75452 mycoplasma
24	30	68.2	783	1	ZF25_HUMAN	Q9uli5 homo sapien
25	29	65.9	174	1	GSPB_KLEPN	P20725 klebsiella
26	29	65.9	259	1	PYRK_METMA	Q8pw55 methanosarc
27	29	65.9	318	1	UL31_EBV	P03183 epstein-bar
28	29	65.9	375	1	BCA2_YEAST	P47176 saccharomyc
29	29	65.9	377	1	CISY_SULSO	P80148 sulfolobus
30	29	65.9	393	1	PELA_ERWCH	P29155 erwinia chr
31	29	65.9	400	1	OPN3_MOUSE	Q9wuk7 mus musculu
32	29	65.9	401	1	RENS_MOUSE	P00796 mus musculu
33	29	65.9	402	1	RENI_MOUSE	P06281 mus musculu
34	29	65.9	404	1	GSPF_ERWCH	P31704 erwinia chr
35	29	65.9	478	1	Y4RH_RHISN	P55641 rhizobium s
36	29	65.9	508	1	YBY9_YEAST	P38109 saccharomyc
37	29	65.9	559	1	SYQ_PASMU	P57847 pasteurella
38	29	65.9	610	1	MET7_SCHPO	O74314 schizosacch
39	29	65.9	663	1	UVRB_AQUAE	O67708 aquifex aeo
40	29	65.9	720	1	Z408_HUMAN	Q9h9d4 homo sapien
41	29	65.9	730	1	CYG4_RAT	Q9wvi4 rattus norv
42	29	65.9	732	1	CYG4_HUMAN	P33402 homo sapien
43	29	65.9	772	1	CPTM_HUMAN	Q92523 homo sapien
44	29	65.9	772	1	CPTM_RAT	Q63704 rattus norv
45	29	65.9	794	1	YG2A_YEAST	P53243 saccharomyc

ALIGNMENTS

RESULT 1
C3L1_HUMAN
ID C3L1_HUMAN STANDARD; PRT; 383 AA.
AC P36222; P30923;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Chitinase-3 like protein 1 precursor (Cartilage glycoprotein-39) (GP-
DE 39) (39 kDa synovial protein) (YKL-40).
GN CHI3L1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Cartilage;
RX MEDLINE=94064658; PubMed=8245017;
RA Hakala B.E., White C., Recklies A.D.;
RT "Human cartilage gp-39, a major secretory product of articular
RT chondrocytes and synovial cells, is a mammalian member of a chitinase
RT protein family.";
RL J. Biol. Chem. 268:25803-25810(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=97386591; PubMed=9244440;
RA Rehli M., Krause S.W., Andreesen R.;
RT "Molecular characterization of the gene for human cartilage gp-39
RT (CHI3L1), a member of the chitinase protein family and marker for
RT late stages of macrophage differentiation.";
RL Genomics 43:221-225(1997).
RN [3]
RP SEQUENCE OF 22-45.
RX MEDLINE=90328983; PubMed=2375755;
RA Nyirkos P., Golds E.E.;
RT "Human synovial cells secrete a 39 kDa protein similar to a bovine
RT mammary protein expressed during the non-lactating period.";
RL Biochem. J. 269:265-268(1990).
CC -!- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN THE CAPACITY OF CELLS TO
CC RESPOND TO AND COPE WITH CHANGES IN THEIR ENVIRONMENT.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- TISSUE SPECIFICITY: PRESENT IN ARTICULAR CHONDROCYTES, SYNOVIAL
CC CELLS AS WELL AS IN LIVER. UNDETECTABLE IN MUSCLE TISSUES, LUNG,
CC PANCREAS, MONONUCLEAR CELLS, OR FIBROBLASTS.
CC -!- PTM: GLYCOSYLATED.
CC -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.
CC ----------

CC ----------

```
DR   EMBL; M80927; AAA16074.1; -.
DR   EMBL; Y08374; CAA69661.1; -.
DR   EMBL; Y08375; CAA69661.1; JOINED.
DR   EMBL; Y08376; CAA69661.1; JOINED.
DR   EMBL; Y08377; CAA69661.1; JOINED.
DR   EMBL; Y08378; CAA69661.1; JOINED.
DR   PIR; A49562; A49562.
DR   PDB; 1LA7; 10-APR-02.
DR   Genew; HGNC:1932; CHI3L1.
DR   MIM; 601525; -.
DR   GO; GO:0005578; C:extracellular matrix; TAS.
DR   GO; GO:0005615; C:extracellular space; TAS.
DR   GO; GO:0005207; F:extracellular matrix glycoprotein; TAS.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   InterPro; IPR001579; Glyco_hydro_18/2.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   ProDom; PD000471; Glyco_hydro_18; 1.
DR   SMART; SM00636; Glyco_18; 1.
DR   PROSITE; PS01095; CHITINASE_18; FALSE_NEG.
KW   Glycoprotein; Signal; 3D-structure.
FT   SIGNAL        1     21
FT   CHAIN        22    383       CHITINASE-3 LIKE PROTEIN 1.
FT   CARBOHYD     60     60       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   383 AA;  42613 MW;  76ADD8298EEEC2D1 CRC64;

  Query Match             100.0%;  Score 44;  DB 1;  Length 383;
  Best Local Similarity   100.0%;  Pred. No. 0.051;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FGRSFTLAS 9
              |||||||||
Db        261 FGRSFTLAS 269


RESULT 2
C3L1_MOUSE
ID   C3L1_MOUSE     STANDARD;      PRT;   381 AA.
AC   Q61362;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Chitinase-3 like protein 1 precursor (Cartilage glycoprotein-39)
DE   (GP-39) (BRP39 protein).
GN   CHI3L1 OR BRP39.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FVB/N; TISSUE=Breast;
RX   MEDLINE=95060797; PubMed=7970700;
RA   Morrison B.W., Leder P.;
RT   "neu and ras initiate murine mammary tumors that share genetic markers
RT   generally absent in c-myc and int-2-initiated tumors.";
RL   Oncogene 9:3417-3426(1994).
CC   -!- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN THE CAPACITY OF CELLS TO
CC       RESPOND TO AND COPE WITH CHANGES IN THEIR ENVIRONMENT (BY
CC       SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Extracellular (By similarity).
CC   -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.
CC   ---------------------------------------------------------------------- ----------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X93035; CAA63603.1; -.
DR   PIR; S61551; S61551.
DR   MGD; MGI:1340899; Chi3l1.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   InterPro; IPR001579; Glyco_hydro_18/2.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   ProDom; PD000471; Glyco_hydro_18; 1.
DR   SMART; SM00636; Glyco_18; 1.
DR   PROSITE; PS01095; CHITINASE_18; FALSE_NEG.
KW   Glycoprotein; Signal.
FT   SIGNAL        1     21       POTENTIAL.
FT   CHAIN        22    381       CHITINASE-3 LIKE PROTEIN 1.
FT   CARBOHYD     60     60       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   381 AA;  43001 MW;  EF6581E8184F045C CRC64;

  Query Match             93.2%;  Score 41;  DB 1;  Length 381;
  Best Local Similarity   88.9%;  Pred. No. 0.23;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FGRSFTLAS 9
              |||:||||
Db        262 FGKSFTLAS 270


RESULT 3
L2CC_DROME
ID   L2CC_DROME     STANDARD;      PRT;   293 AA.
AC   P24156;
DT   01-MAR-1992 (Rel. 21, Created)
DT   01-MAR-1992 (Rel. 21, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   L(2)37CC protein.
GN   L(2)37CC.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86312887; PubMed=3092183;
RA   Eveleth D.D. Jr., Marsh J.L.;
RT   "Sequence and expression of the Cc gene, a member of the dopa
RT   decarboxylase gene cluster of Drosophila: possible translational
RT   regulation.";
RL   Nucleic Acids Res. 14:6169-6183(1986).
CC   -!- FUNCTION: REQUIRED FOR LARVAL METABOLISM OR FOR THE PROGRESSION
CC       OF THE LARVA INTO A PUPA.
CC   -!- SIMILARITY: BELONGS TO THE PROHIBITIN FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X04228; CAA27810.1; -.
DR   EMBL; X04227; CAA27807.1; -.
DR   PIR; C25511; C25511.
DR   FlyBase; FBgn0002031; l(2)37Cc.
DR   InterPro; IPR001107; Band_7.
DR   InterPro; IPR000163; Prohibitin.
DR   Pfam; PF01145; Band_7; 1.
DR   PRINTS; PR00679; PROHIBITIN.
DR   SMART; SM00244; PHB; 1.
KW   Developmental protein.
SQ   SEQUENCE   203 AA;  22817 MW;  B15D085CC0862A11 CRC64;

  Query Match             81.8%;  Score 36;  DB 1;  Length 203;
  Best Local Similarity   87.5%;  Pred. No. 1.4;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FGRSFTLA 8
              ||| ||||
```

```
Db         98 FGREFTLA 105


RESULT 4
C3L2_HUMAN
ID   C3L2_HUMAN     STANDARD;      PRT:   390 AA.
AC   Q15782; Q15749; Q15783;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Chitinase 3-like protein 2 precursor (YKL-39) (Chondrocyte protein
DE   39).
GN   CHI3L2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS LONG AND SHORT).
RA   Grossman A., Matsuyama T., Baker E., Waterhouse P., Sutherland G.R.,
RA   Mak T.W.;
RT   "Cloning of a novel lymphoid restricted human chitinase and
RT   localization to 1p13.3.";
RL   Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM SHORT).
RC   TISSUE=Articular cartilage;
RX   MEDLINE=96325055; PubMed=8702629;
RA   Hu B., Trinh K., Figueira W.F., Price P.A.;
RT   "Isolation and sequence of a novel human chondrocyte protein related
RT   to mammalian members of the chitinase protein family.";
RL   J. Biol. Chem. 271:19415-19420(1996).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=Short;
CC         IsoId=Q15782-1; Sequence=Displayed;
CC       Name=Long;
CC         IsoId=Q15782-2; Sequence=VSP_001066;
CC         Note=No experimental confirmation available;
CC   -!- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN CHONDROCYTES, FOLLOWED
CC       BY SYNOVIOCYTES, LUNG AND HEART. NOT DETECTED IN BRAIN, SPLEEN,
CC       PANCREAS, AND LIVER.
CC   -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U58515; AAB04534.1; -.
DR   EMBL; U58514; AAB04533.1; -.
DR   EMBL; U49835; AAC50597.1; ALT_INIT.
DR   Genew; HGNC:1933; CHI3L2.
DR   MIM; 601526; -.
DR   GO; GO:0005615; C:extracellular space; TAS.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   InterPro; IPR001579; Glyco_hydro_18/2.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   ProDom; PD000471; Glyco_hydro_18; 1.
DR   SMART; SM00636; Glyco_18; 1.
DR   PROSITE; PS01095; CHITINASE_18; FALSE_NEG.
KW   Hydrolase; Signal; Alternative splicing.
FT   SIGNAL       1     26       POTENTIAL.
FT   CHAIN       27    390       CHITINASE 3-LIKE PROTEIN 2.
FT   CARBOHYD    35     35       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC     1     12       MGATTMDQKSLW -> SQQLWLGSPDEVWLYLVCEHTHIFT
FT                               AIIWDSRTRFGSTDPSFHPSI (in isoform Long).
FT                               /FTId=VSP_001066.
SQ   SEQUENCE   390 AA;  43500 MW;  97B86A2F3AA35677 CRC64;


  Query Match             81.8%;  Score 36;  DB 1;  Length 390;
  Best Local Similarity   77.8%;  Pred. No. 2.9;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FGRSFTLAS 9
              :| ||||||
Db        269 YGHSFTLAS 277


RESULT 5
CHIT_MANSE
ID   CHIT_MANSE     STANDARD;      PRT:   554 AA.
AC   P36362;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-FEB-1995 (Rel. 31, Last annotation update)
DE   Endochitinase precursor (EC 3.2.1.14).
OS   Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Sphingiodea;
OC   Sphingidae; Sphinginae; Manduca.
OX   NCBI_TaxID=7130;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93357793; PubMed=8353525;
RA   Kramer K.J., Corpuz L., Choi H.K., Muthukrishnan S.;
RT   "Sequence of a cDNA and expression of the gene encoding epidermal and
RT   gut chitinases of Manduca sexta.";
RL   Insect Biochem. Mol. Biol. 23:691-701(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97215580; PubMed=9061927;
RA   Choi H.K., Choi K.H., Kramer K.J., Muthukrishnan S.;
RT   "Isolation and characterization of a genomic clone for the gene of an
RT   insect molting enzyme, chitinase.";
RL   Insect Biochem. Mol. Biol. 27:37-47(1997).
CC   -!- FUNCTION: DIGEST CHITIN IN THE EXOSKELETON DURING THE MOLTING
CC       PROCESS.
CC   -!- CATALYTIC ACTIVITY: Hydrolysis of the 1,4-beta-linkages of N-
CC       acetyl-D-glucosamine polymers of chitin.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: EPIDERMIS AND GUT.
CC   -!- DEVELOPMENTAL STAGE: HIGH LEVELS SEEN IN THE EPIDERMIS ON DAY 0,
CC       BUT RAPIDLY DISAPPEARS AND IS UNDETECTED ON DAYS 1-4 OF FIFTH
CC       INSTAR. IT REAPPEARS ON DAY 5 AND PEAKS ON DAY 7 AFTER WHICH A
CC       RAPID DECLINE IS SEEN. IN THE GUT IS DETECTED ON DAY 6 WITH LOWER
CC       LEVELS SEEN ON DAYS 0, 7 AND 8.
CC   -!- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
CC       HYDROLASES).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U02270; AAC04924.1; -.
DR   EMBL; L49234; AAB53952.1; -.
DR   PIR; A56596; A56596.
DR   InterPro; IPR002557; Chitin_bind_PerA.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   InterPro; IPR001579; Glyco_hydro_18/2.
DR   Pfam; PF01607; CBM_14; 1.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   ProDom; PD000471; Glyco_hydro_18; 1.
DR   SMART; SM00494; ChtBD2; 1.
DR   SMART; SM00636; Glyco_18; 1.
DR   PROSITE; PS01095; CHITINASE_18; 1.
KW   Hydrolase; Glycosidase; Chitin degradation; Signal; Glycoprotein.
FT   SIGNAL       1     19       POTENTIAL.
FT   CHAIN       20    554       ENDOCHITINASE.
```

```
FT   DOMAIN      396    453       SER/THR-RICH.
FT   ACT_SITE    146    146       PROTON DONOR (BY SIMILARITY).
FT   CARBOHYD     85     85       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    303    303       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    407    407       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    545    545       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   554 AA;  62203 MW;  3989D756C96CD490 CRC64;

  Query Match             79.5%;  Score 35;  DB 1;  Length 554;
  Best Local Similarity   66.7%;  Pred. No. 7;
  Matches    6;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FGRSFTLAS 9
              :|||||::
Db        270 YGRSFTLSA 278


RESULT 6
ZBT1_HUMAN
ID   ZBT1_HUMAN     STANDARD;      PRT;   713 AA.
AC   Q9Y2K1;
DT   15-SEP-2003 (Rel. 42, Created)
DT   15-SEP-2003 (Rel. 42, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Zinc finger and BTB domain containing protein 1.
GN   ZBTB1 OR KIAA0997.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Eye;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE OF 1-644 FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=99246063; PubMed=10231032;
RA   Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirosawa M.,
RA   Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT   "Prediction of the coding sequences of unidentified human genes. XIII.
RT   The complete sequences of 100 new cDNA clones from brain which code
RT   for large proteins in vitro.";
RL   DNA Res. 6:63-70(1999).
CC   -!- FUNCTION: May be involved in transcriptional regulation.
CC   -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC   -!- SIMILARITY: Contains 1 BTB/POZ domain.
CC   -!- SIMILARITY: Contains 8 C2H2-type zinc fingers.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed.  Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; BC050719; AAH50719.1; -.
DR   EMBL; AB023214; BAA76841.1; -.
DR   Genew; HGNC:20259; ZBTB1.
DR   InterPro; IPR000210; BTB_POZ.
DR   InterPro; IPR007087; Znf_C2H2.
DR   Pfam; PF00651; BTB; 1.
DR   Pfam; PF00096; zf-C2H2; 5.
DR   SMART; SM00225; BTB; 1.
DR   SMART; SM00355; ZnF_C2H2; 8.
DR   PROSITE; PS50097; BTB; 1.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2_1; 4.
DR   PROSITE; PS50157; ZINC_FINGER_C2H2_2; 3.
KW   Transcription regulation; DNA-binding; Zinc-finger; Metal-binding;
KW   Nuclear protein; Repeat.
FT   DOMAIN       24     91       BTB.
FT   ZN_FING     216    238       C2H2-TYPE 1 (ATYPICAL).
FT   ZN_FING     448    470       C2H2-TYPE 2 (ATYPICAL).
FT   ZN_FING     534    556       C2H2-TYPE 3 (ATYPICAL).
FT   ZN_FING     578    600       C2H2-TYPE 4.
FT   ZN_FING     606    628       C2H2-TYPE 5.
FT   ZN_FING     634    656       C2H2-TYPE 6.
FT   ZN_FING     662    684       C2H2-TYPE 7.
FT   ZN_FING     686    709       C2H2-TYPE 8.
FT   CONFLICT    203    203       T -> N (IN REF. 2).
FT   CONFLICT    334    334       T -> I (IN REF. 2).
FT   CONFLICT    633    644       RYVCSICDQGNF -> SGEIGPSKPVEK (IN REF. 2).
SQ   SEQUENCE   713 AA;  82003 MW;  712DFEE75C9DB302 CRC64;

  Query Match             75.0%;  Score 33;  DB 1;  Length 713;
  Best Local Similarity   77.8%;  Pred. No. 25;
  Matches    7;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 FGRSFTLAS 9
              |||||| |
Db        212 FGRSFTCDS 220


RESULT 7
ZBT1_MOUSE
ID   ZBT1_MOUSE     STANDARD;      PRT;   713 AA.
AC   Q91VL9; Q8CDP7;
DT   15-SEP-2003 (Rel. 42, Created)
DT   15-SEP-2003 (Rel. 42, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Zinc finger and BTB domain containing protein 1.
GN   ZBTB1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Embryo;
RX   MEDLINE=22354683; PubMed=12466851;
RA   Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA   Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA   Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA   Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA   Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA   Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA   Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA   Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA   Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA   Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA   Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA   Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA   Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA   Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA   Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
```

```
RA   Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA   Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA   Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA   Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA   Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA   Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA   Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA   Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA   Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA   Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA   Birney E., Hayashizaki Y.;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Breast tumor;
RX   MEDLINE=22388287; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- FUNCTION: May be involved in transcriptional regulation.
CC   -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC   -!- SIMILARITY: Contains 1 BTB/POZ domain.
CC   -!- SIMILARITY: Contains 8 C2H2-type zinc fingers.
CC   ---------------------------------------------------------------------- ----------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AK029762; BAC26603.1; -.
DR   EMBL; AK035036; BAC28921.1; -.
DR   EMBL; AK049397; BAC33733.1; -.
DR   EMBL; BC012239; AAH12239.1; -.
DR   MGD; MGI:2442326; C430003J21Rik.
DR   InterPro; IPR000210; BTB_POZ.
DR   InterPro; IPR007087; Znf_C2H2.
DR   Pfam; PF00651; BTB; 1.
DR   Pfam; PF00096; zf-C2H2; 8.
DR   SMART; SM00225; BTB; 1.
DR   SMART; SM00355; ZnF_C2H2; 8.
DR   PROSITE; PS50097; BTB; 1.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2_1; 4.
DR   PROSITE; PS50157; ZINC_FINGER_C2H2_2; 4.
KW   Transcription regulation; DNA-binding; Zinc-finger; Metal-binding;
KW   Nuclear protein; Repeat.
FT   DOMAIN      24     91       BTB.
FT   ZN_FING     216    238      C2H2-TYPE 1 (ATYPICAL).
FT   ZN_FING     448    470      C2H2-TYPE 2 (ATYPICAL).
FT   ZN_FING     534    556      C2H2-TYPE 3 (ATYPICAL).
FT   ZN_FING     578    600      C2H2-TYPE 4.
FT   ZN_FING     606    628      C2H2-TYPE 5.
FT   ZN_FING     634    656      C2H2-TYPE 6.
FT   ZN_FING     662    684      C2H2-TYPE 7.
FT   ZN_FING     686    709      C2H2-TYPE 8.
FT   CONFLICT    643    643      N -> K (IN REF. 1; BAC26603).
SQ   SEQUENCE   713 AA;  81951 MW;  7ECD7EA1E9C88500 CRC64;

  Query Match             75.0%;  Score 33;  DB 1;  Length 713;
  Best Local Similarity   77.8%;  Pred. No. 25;
  Matches    7;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 FGRSFTLAS 9
              ||||||  |
Db        212 FGRSFTCDS 220


RESULT 8
PYRK_METAC
ID   PYRK_METAC     STANDARD;      PRT:   259 AA.
AC   P58886;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Probable dihydroorotate dehydrogenase electron transfer subunit.
GN   PYRK OR MA0584.
OS   Methanosarcina acetivorans.
OC   Archaea; Euryarchaeota; Euryarchaeota orders incertae sedis;
OC   Methanosarcinales; Methanosarcinaceae; Methanosarcina.
OX   NCBI_TaxID=2214;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C2A / ATCC 35395 / DSM 2834;
RX   MEDLINE=21929759; PubMed=11932238;
RA   Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,
RA   FitzHugh W., Calvo S., Engels R., Smirnov S., Atnoor D., Brown A.,
RA   Allen N., Naylor J., Stange-Thomann N., DeArellano K., Johnson R.,
RA   Linton L., McEwan P., McKernan K., Talamas J., Tirrell A., Ye W.,
RA   Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,
RA   Hedderich R., Ingram-Smith C., Kuettner H.C., Krzycki J.A.,
RA   Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
RA   Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,
RA   Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
RA   Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
RA   Metcalf W.W., Birren B.;
RT   "The genome of Methanosarcina acetivorans reveals extensive metabolic
RT   and physiological diversity.";
RL   Genome Res. 12:532-542(2002).
CC   -!- FUNCTION: Is responsible for channelling the electrons from the
CC       oxidation of dihydroorotate from the FMN redox center in the pyrD
CC       subunit to the ultimate electron acceptor NAD(+) (By similarity).
CC   -!- COFACTOR: Binds a 2Fe-2S cluster and FAD (By similarity).
CC   -!- PATHWAY: Pyrimidine biosynthesis; fourth step.
CC   -!- SUBUNIT: Heterotetramer of 2 pyrK and 2 pyrD subunits (By
CC       similarity).
CC   -!- SIMILARITY: Belongs to the pyrK family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AE010719; AAM04028.1; -.
DR   HAMAP; MF_01211; -; 1.
DR   InterPro; IPR006058; 2Fe2S_ferredoxin.
DR   InterPro; IPR001834; Cyt_B5_reductase.
DR   InterPro; IPR001433; Oxred_FAD/NAD(P).
DR   Pfam; PF00970; FAD_binding_6; 1.
DR   Pfam; PF00175; NAD_binding_1; 1.
DR   PROSITE; PS00197; 2FE2S_FERREDOXIN; 1.
KW   Pyrimidine biosynthesis; Transport; Electron transport; Metal-binding;
KW   Iron; Iron-sulfur; 2Fe-2S; Flavoprotein; FAD; Complete proteome.
```

```
FT   METAL       211    211       IRON-SULFUR 1 (2FE-2S) (BY SIMILARITY).
FT   METAL       216    216       IRON-SULFUR 1 (2FE-2S) (BY SIMILARITY).
FT   METAL       219    219       IRON-SULFUR 2 (2FE-2S) (BY SIMILARITY).
FT   METAL       229    229       IRON-SULFUR 2 (2FE-2S) (BY SIMILARITY).
SQ   SEQUENCE   259 AA;  27982 MW;  85E107DCF22A912E CRC64;

  Query Match             72.7%;  Score 32;  DB 1;  Length 259;
  Best Local Similarity   66.7%;  Pred. No. 14;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 FGRSFTLAS 9
              ||: ||| |
Db         83 FGKGFTLPS 91


RESULT 9
BCHX_RHOCA
ID   BCHX_RHOCA     STANDARD;      PRT;   333 AA.
AC   P26177;
DT   01-MAY-1992 (Rel. 22, Created)
DT   01-MAY-1992 (Rel. 22, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Chlorophyllide reductase 35.5 kDa chain (EC 1.3.1.-) (Chlorin
DE   reductase).
GN   BCHX.
OS   Rhodobacter capsulatus (Rhodopseudomonas capsulata).
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;
OC   Rhodobacteraceae; Rhodobacter.
OX   NCBI_TaxID=1061;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SB1003 / St Louis;
RX   MEDLINE=93224464; PubMed=8468299;
RA   Burke D.H., Alberti M., Hearst J.E.;
RT   "The Rhodobacter capsulatus chlorin reductase-encoding locus, bchA,
RT   consists of three genes, bchX, bchY, and bchZ.";
RL   J. Bacteriol. 175:2407-2413(1993).
CC   -!- FUNCTION: CONVERTS CHLOROPHYLLS (CHL) INTO BACTERIOCHLOROPHYLLS
CC       (BCHL) BY REDUCING RING B OF THE TETRAPYRROLE.
CC   -!- PATHWAY: Bacteriochlorophyll biosynthesis.
CC   -!- SIMILARITY: BELONGS TO THE NIFH / BCHL / CHLL FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; Z11165; CAA77548.1; -.
DR   PIR; S17823; S17823.
DR   HSSP; P00456; 1CP2.
DR   InterPro; IPR000392; NitrogenaseII.
DR   Pfam; PF00142; fer4_NifH; 1.
DR   PRINTS; PR00091; NITROGNASEII.
DR   PROSITE; PS00692; NIFH_FRXC_2; 1.
DR   PROSITE; PS00746; NIFH_FRXC_1; 1.
KW   Oxidoreductase; Photosynthesis; Bacteriochlorophyll biosynthesis;
KW   ATP-binding; Iron-sulfur; 4Fe-4S.
FT   NP_BIND      42     49       ATP (POTENTIAL).
FT   METAL       130    130       IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT   METAL       165    165       IRON-SULFUR (4FE-4S) (BY SIMILARITY).
SQ   SEQUENCE   333 AA;  35573 MW;  CD697D641C8943EB CRC64;

  Query Match             72.7%;  Score 32;  DB 1;  Length 333;
  Best Local Similarity   75.0%;  Pred. No. 18;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 GRSFTLAS 9
              |:|||||:
Db         47 GKSFTLAN 54


RESULT 10
BCHX_RHOSH
ID   BCHX_RHOSH     STANDARD;      PRT;   333 AA.
AC   Q02431; Q9RFC2;
DT   01-JUL-1993 (Rel. 26, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Chlorophyllide reductase 35.5 kDa chain (EC 1.3.1.-) (Chlorin
DE   reductase).
GN   BCHX.
OS   Rhodobacter sphaeroides (Rhodopseudomonas sphaeroides).
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;
OC   Rhodobacteraceae; Rhodobacter.
OX   NCBI_TaxID=1063;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 17023 / 2.4.1 / NCIB 8253 / DSM 158;
RX   MEDLINE=93173096; PubMed=8437569;
RA   McGlynn P., Hunter C.N.;
RT   "Genetic analysis of the bchC and bchA genes of Rhodobacter
RT   sphaeroides.";
RL   Mol. Gen. Genet. 236:227-234(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 17023 / 2.4.1 / NCIB 8253 / DSM 158;
RX   MEDLINE=20115911; PubMed=10648776;
RA   Choudhary M., Kaplan S.;
RT   "DNA sequence analysis of the photosynthesis region of Rhodobacter
RT   sphaeroides 2.4.1.";
RL   Nucleic Acids Res. 28:862-867(2000).
CC   -!- FUNCTION: CONVERTS CHLOROPHYLLS (CHL) INTO BACTERIOCHLOROPHYLLS
CC       (BCHL) BY REDUCING RING B OF THE TETRAPYRROLE.
CC   -!- PATHWAY: Bacteriochlorophyll biosynthesis.
CC   -!- SIMILARITY: BELONGS TO THE NIFH / BCHL / CHLL FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AJ010302; CAB38747.1; -.
DR   EMBL; AF195122; AAF24297.1; -.
DR   PIR; S30915; S30915.
DR   PIR; T50753; T50753.
DR   HSSP; P00456; 1CP2.
DR   InterPro; IPR000392; NitrogenaseII.
DR   Pfam; PF00142; fer4_NifH; 1.
DR   PRINTS; PR00091; NITROGNASEII.
DR   PROSITE; PS00692; NIFH_FRXC_2; 1.
DR   PROSITE; PS00746; NIFH_FRXC_1; 1.
KW   Oxidoreductase; Photosynthesis; Bacteriochlorophyll biosynthesis;
KW   ATP-binding; Iron-sulfur; 4Fe-4S.
FT   NP_BIND      42     49       ATP (BY SIMILARITY).
FT   METAL       130    130       IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT   METAL       165    165       IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT   CONFLICT    127    127       A -> G (IN REF. 1).
SQ   SEQUENCE   333 AA;  35589 MW;  F8E9AD672AD8A501 CRC64;

  Query Match             72.7%;  Score 32;  DB 1;  Length 333;
  Best Local Similarity   75.0%;  Pred. No. 18;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 GRSFTLAS 9
              |:|||||:
Db         47 GKSFTLAN 54


RESULT 11
IRT_CAEEL
```

```
ID   IRT_CAEEL      STANDARD;      PRT;   405 AA.
AC   Q27371;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Troponin T.
GN   MUP-2 OR T22E5.5.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2; TISSUE=Muscle;
RX   MEDLINE=96175302; PubMed=8601585;
RA   Myers C.D., Goh P.-Y., Allen T.S., Bucher E.A., Bogaert T.;
RT   "Developmental genetic analysis of troponin T mutations in striated
RT   and nonstriated muscle cells of Caenorhabditis elegans.";
RL   J. Cell Biol. 132:1061-1077(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Minx P.;
RL   Submitted (JAN-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: TROPONIN T IS THE TROPOMYOSIN-BINDING SUBUNIT OF
CC       TROPONIN, THE THIN FILAMENT REGULATORY COMPLEX WHICH CONFERS
CC       CALCIUM-SENSITIVITY TO STRIATED MUSCLE ACTOMYOSIN ATPASE ACTIVITY
CC       (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE TROPONIN T FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U44759; AAA86906.1; -.
DR   EMBL; U43282; AAA83615.1; -.
DR   PIR; T16922; T16922.
DR   WormPep; T22E5.5; CE04994.
DR   InterPro; IPR001978; Troponin.
DR   Pfam; PF00992; Troponin; 1.
FT   DOMAIN       11     23       POLY-GLU.
FT   DOMAIN       75     78       POLY-GLU.
SQ   SEQUENCE   405 AA;  47041 MW;  1D3AC9EEB641C65B CRC64;

  Query Match             72.7%;  Score 32;  DB 1;  Length 405;
  Best Local Similarity   75.0%;  Pred. No. 22;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 GRSFTLAS 9
              ||:|||:||
Db        174 GRNFTVAS 181


RESULT 12
ZN12_MICSA
ID   ZN12_MICSA     STANDARD;      PRT;   415 AA.
AC   P38621;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Zinc finger protein ZFMSA12A.
OS   Micropterus salmoides (Largemouth bass).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidei;
OC   Centrarchidae; Micropterus.
OX   NCBI_TaxID=27706;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Ovary;
RX   MEDLINE=93382800; PubMed=8371993;
RA   Ogilvie M.K., Smith J.F., Hanas J.S.;
RT   "Sequence divergence of a TFIIIA-type zinc finger protein from fish
RT   ovarian tissue.";
RL   Nucleic Acids Res. 21:4152-4152(1993).
CC   -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z22756; CAA80438.1; -.
DR   PIR; S34140; S34140.
DR   HSSP; P08047; 1SP2.
DR   InterPro; IPR007087; Znf_C2H2.
DR   InterPro; IPR007086; Znf_C2H2_sub.
DR   Pfam; PF00096; zf-C2H2; 11.
DR   PRINTS; PR00048; ZINCFINGER.
DR   ProDom; PD000003; Znf_C2H2; 4.
DR   SMART; SM00355; ZnF_C2H2; 12.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2_1; 11.
DR   PROSITE; PS50157; ZINC_FINGER_C2H2_2; 12.
KW   Zinc-finger; Metal-binding; DNA-binding; Repeat; Nuclear protein.
FT   ZN_FING      78    100      C2H2-TYPE.
FT   ZN_FING     106    129      C2H2-TYPE.
FT   ZN_FING     134    156      C2H2-TYPE.
FT   ZN_FING     161    183      C2H2-TYPE.
FT   ZN_FING     189    211      C2H2-TYPE.
FT   ZN_FING     217    239      C2H2-TYPE.
FT   ZN_FING     245    267      C2H2-TYPE.
FT   ZN_FING     273    295      C2H2-TYPE.
FT   ZN_FING     301    323      C2H2-TYPE.
FT   ZN_FING     329    351      C2H2-TYPE.
FT   ZN_FING     357    379      C2H2-TYPE.
FT   ZN_FING     385    407      C2H2-TYPE.
SQ   SEQUENCE   415 AA;  47237 MW;  3A2D2B9CC40278B8 CRC64;

  Query Match             72.7%;  Score 32;  DB 1;  Length 415;
  Best Local Similarity   75.0%;  Pred. No. 23;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 GRSFTLAS 9
              ||:|||:||
Db        335 GRGFTMAS 342


RESULT 13
FAT1_SCHPO
ID   FAT1_SCHPO     STANDARD;      PRT;  1385 AA.
AC   O13735; Q9UTJ0;
DT   15-JUL-1998 (Rel. 36, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Actin interacting protein 3 homolog.
GN   FAT1 OR SPAC15A10.16 OR SPAC15E1.01.
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
OC   Schizosaccharomyces.
OX   NCBI_TaxID=4896;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972;
RX   MEDLINE=21848401; PubMed=11859360;
RA   Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA   Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA   Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA   Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
```

```
RA   Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA   Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA   James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA   Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA   Oliver K., O'Neil S., Pearson D., Quail M.A., Rabbinowitsch E.,
RA   Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA   Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA   Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA   Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA   Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA   Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,
RA   Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA   Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA   Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA   Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA   Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA   Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA   Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA   Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA   Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT   "The genome sequence of Schizosaccharomyces pombe.";
RL   Nature 415:871-880(2002).
RN   [2]
RP   IDENTIFICATION, AND GENE NAME.
RX   MEDLINE=20143585; PubMed=10679021;
RA   Jin H., Amberg D.C.;
RT   "The secretory pathway mediates localization of the cell polarity
RT   regulator Aip3p/Bud6p.";
RL   Mol. Biol. Cell 11:647-661(2000).
CC   -!- FUNCTION: INVOLVED IN THE ORGANIZATION AND/OR FUNCTION OF THE
CC       ACTIN CYTOSKELETON.
CC   -!- SIMILARITY: TO YEAST BUD6.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z97208; CAB10112.1; -.
DR   EMBL; AL109770; CAB52420.1; -.
DR   GeneDB_SPombe; SPAC15A10.16; -.
DR   InterPro; IPR005613; AIP3.
DR   Pfam; PF03915; AIP3; 1.
KW   Coiled coil; Cytoskeleton.
FT   DOMAIN     1009    1096       COILED COIL (POTENTIAL).
SQ   SEQUENCE   1385 AA;  154325 MW;  ABB3D40CC4FF7537 CRC64;

  Query Match           72.7%;  Score 32;  DB 1;  Length 1385;
  Best Local Similarity   66.7%;  Pred. No. 87;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 FGRSFTLAS 9
              ||| | |:|
Db       1005 FGRGFVLSS 1013


RESULT 14
R40C_HUMAN
ID   R40C_HUMAN     STANDARD;      PRT;   281 AA.
AC   Q96S21; O60795;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Ras-related protein Rab-40C (SOCS box containing protein RAR3)
DE   (Rar-like protein).
GN   RAB40C OR RARL.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21096910; PubMed=11157797;
RA   Daniels R.J., Peden J.F., Lloyd C., Horsley S.W., Clark K.,
RA   Tufarelli C., Kearney L., Buckle V.J., Doggett N.A., Flint J.,
RA   Higgs D.R.;
RT   "Sequence, structure and pathology of the fully annotated terminal 2
RT   Mb of the short arm of human chromosome 16.";
RL   Hum. Mol. Genet. 10:339-352(2001).
RN   [2]
RP   SEQUENCE OF 46-281 FROM N.A.
RA   Percy C.;
RL   Submitted (OCT-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO THE SMALL GTPASE SUPERFAMILY. RAB FAMILY.
CC   -!- SIMILARITY: Contains 1 SOCS box domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE006464; AAK61236.1; -.
DR   EMBL; Z84479; CAB06484.1; -.
DR   HSSP; P05713; 3RAB.
DR   Genew; HGNC:18285; RAB40C.
DR   InterPro; IPR003579; GTPase_Rab.
DR   InterPro; IPR001806; Ras_trasfrmng.
DR   InterPro; IPR005225; Small_GTP.
DR   InterPro; IPR001496; SOCS.
DR   Pfam; PF00071; ras; 1.
DR   PRINTS; PR00449; RASTRNSFRMNG.
DR   SMART; SM00175; RAB; 1.
DR   SMART; SM00253; SOCS; 1.
DR   TIGRFAMs; TIGR00231; small_GTP; 1.
DR   PROSITE; PS50225; SOCS; 1.
KW   GTP-binding; Lipoprotein; Prenylation.
FT   NP_BIND      21     28       GTP (BY SIMILARITY).
FT   NP_BIND      69     73       GTP (BY SIMILARITY).
FT   NP_BIND     126    129       GTP (BY SIMILARITY).
FT   DOMAIN      175    228       SOCS BOX.
FT   LIPID       278    278       GERANYL-GERANYL (BY SIMILARITY).
FT   CONFLICT     46     47       SN -> LA (IN REF. 2).
SQ   SEQUENCE   281 AA;  31304 MW;  478F974645AA91EC CRC64;

  Query Match           70.5%;  Score 31;  DB 1;  Length 281;
  Best Local Similarity   75.0%;  Pred. No. 25;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 GRSFTLAS 9
              |||::|||
Db        238 GRSYSLAS 245


RESULT 15
R40C_MOUSE
ID   R40C_MOUSE     STANDARD;      PRT;   281 AA.
AC   Q8VHQ4;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Ras-related protein Rab-40C (SOCS box containing protein RAR3).
GN   RAB40C.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6;
RA   Kile B.T., Hilton D.J., Nicola N.A.;
```

```
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Eye;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- SIMILARITY: BELONGS TO THE SMALL GTPASE SUPERFAMILY. RAB FAMILY.
CC   -!- SIMILARITY: Contains 1 SOCS box domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF422144; AAL60515.1; -.
DR   EMBL; BC027008; AAH27008.1; -.
DR   HSSP; P05713; 3RAB.
DR   MGD; MGI:2183454; Rab40c.
DR   InterPro; IPR003579; GTPase_Rab.
DR   InterPro; IPR001806; Ras_trnsfrmng.
DR   InterPro; IPR005225; Small_GTP.
DR   InterPro; IPR001496; SOCS.
DR   Pfam; PF00071; ras; 1.
DR   PRINTS; PR00449; RASTRNSFRMNG.
DR   SMART; SM00175; RAB; 1.
DR   SMART; SM00253; SOCS; 1.
DR   TIGRFAMs; TIGR00231; small_GTP; 1.
DR   PROSITE; PS50225; SOCS; 1.
KW   GTP-binding; Lipoprotein; Prenylation.
FT   NP_BIND      21     28       GTP (BY SIMILARITY).
FT   NP_BIND      69     73       GTP (BY SIMILARITY).
FT   NP_BIND     126    129       GTP (BY SIMILARITY).
FT   DOMAIN      175    228       SOCS BOX.
FT   LIPID       278    278       GERANYL-GERANYL (BY SIMILARITY).
SQ   SEQUENCE   281 AA;  31348 MW;  3625546777BA90A8 CRC64;

  Query Match             70.5%;  Score 31;  DB 1;  Length 281;
  Best Local Similarity   75.0%;  Pred. No. 25;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 GRSFTLAS 9
              !|||:|||
Db        238 GRSYSLAS 245


Search completed: September 24, 2003, 13:42:23
Job time : 6.66292 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 24, 2003, 13:33:45 ; Search time 29.2247 Seconds
(without alignments)
79.469 Million cell updates/sec

Title: US-09-744-282-1
Perfect score: 44
Sequence: 1 FGRSFTLAS 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	44	100.0	383	4	Q96HI7	Q96hi7 homo sapien
2	44	100.0	383	4	Q8IVA4	Q8iva4 homo sapien
3	44	100.0	383	6	Q8SPQ0	Q8spq0 capra hircu
4	44	100.0	383	6	Q29411	Q29411 sus scrofa
5	41	93.2	332	6	O18949	O18949 bos taurus
6	41	93.2	352	11	Q9WTV1	Q9wtv1 rattus norv
7	41	93.2	381	11	Q99J84	Q99j84 mus musculu
8	41	93.2	387	4	Q9H3V8	Q9h3v8 homo sapien
9	41	93.2	389	11	Q8BKL8	Q8bkl8 mus musculu
10	41	93.2	396	11	Q9D7Q1	Q9d7q1 mus musculu
11	41	93.2	466	4	Q13231	Q13231 homo sapien
12	38	86.4	268	5	Q8MSI4	Q8msi4 drosophila
13	38	86.4	431	5	Q8ISH5	Q8ish5 araneus ven
14	38	86.4	460	5	Q9W2M7	Q9w2m7 drosophila
15	38	86.4	525	5	O44079	O44079 anopheles g
16	38	86.4	553	5	P91731	P91731 hyphantria
17	38	86.4	983	5	Q9VZV2	Q9vzv2 drosophila
18	38	86.4	1013	5	Q960M0	Q960m0 drosophila
19	36	81.8	276	5	Q9VIZ4	Q9viz4 drosophila
20	36	81.8	390	4	Q96F97	Q96f97 homo sapien
21	36	81.8	1074	16	Q9X257	Q9x257 thermotoga
22	36	81.8	2838	5	Q8MP05	Q8mp05 tenebrio mo
23	35	79.5	543	5	Q9GR93	Q9gr93 bombyx mori
24	35	79.5	543	5	Q9GV05	Q9gv05 bombyx mori
25	35	79.5	544	5	Q9GQC4	Q9gqc4 bombyx mori
26	35	79.5	552	5	Q9GV44	Q9gv44 spodoptera
27	35	79.5	557	5	Q8MTK0	Q8mtk0 choristoneu
28	35	79.5	565	5	P90710	P90710 bombyx mori
29	35	79.5	565	5	Q9GPG9	Q9gpg9 bombyx mand
30	35	79.5	566	5	Q8WR52	Q8wr52 bombyx mori
31	35	79.5	765	10	Q9SHL3	Q9shl3 arabidopsis
32	35	79.5	929	5	Q8MY79	Q8my79 haemaphysal
33	34	77.3	314	16	Q9KAQ3	Q9kaq3 bacillus ha
34	34	77.3	1030	3	Q9HFQ9	Q9hfq9 emericella
35	34	77.3	1080	5	Q9BL16	Q9bl16 bombyx mori
36	33	75.0	122	16	Q9EX60	Q9ex60 streptomyce
37	33	75.0	249	2	Q9L3B3	Q9l3b3 rhizobium l
38	33	75.0	267	16	Q9PDV7	Q9pdv7 xylella fas
39	33	75.0	405	5	O97403	O97403 phaedon coc
40	33	75.0	644	4	Q9Y2K1	Q9y2k1 homo sapien
41	33	75.0	684	17	O29784	O29784 archaeoglob
42	33	75.0	713	11	Q91VL9	Q91vl9 mus musculu
43	33	75.0	713	11	Q8CDP7	Q8cdp7 mus musculu
44	33	75.0	1635	5	O17412	O17412 aedes aegyp
45	32	72.7	114	17	O59326	O59326 pyrococcus

ALIGNMENTS

RESULT 1
Q96HI7
ID Q96HI7 PRELIMINARY; PRT; 383 AA.
AC Q96HI7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Similar to chitinase 3-like 1 (Cartilage glycoprotein-39).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC008568; AAH08568.1; -.
DR InterPro; IPR001223; Glyco_hydro_18.
DR InterPro; IPR005829; Sug_transporter.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR SMART; SM00636; Glyco_18; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
SQ SEQUENCE 383 AA; 42625 MW; 6C9EF133BDC7C2D1 CRC64;

Query Match 100.0%; Score 44; DB 4; Length 383;
Best Local Similarity 100.0%; Pred. No. 0.39;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FGRSFTLAS 9
|||||||||
Db 261 FGRSFTLAS 269

RESULT 2
Q8IVA4
ID Q8IVA4 PRELIMINARY; PRT; 383 AA.
AC Q8IVA4;

```
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Similar to chitinase 3-like 1 (Cartilage glycoprotein-39).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RA   Strausberg R.;
RL   Submitted (OCT-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC038354; AAH38354.1; -.
SQ   SEQUENCE   383 AA;  42526 MW;  7B345433BDC7C6C9 CRC64;

  Query Match           100.0%;  Score 44;  DB 4;  Length 383;
  Best Local Similarity  100.0%;  Pred. No. 0.39;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FGRSFTLAS 9
              |||||||||
Db        261 FGRSFTLAS 269


RESULT 3
Q8SPQ0
ID   Q8SPQ0      PRELIMINARY;      PRT;   383 AA.
AC   Q8SPQ0;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   BP40 precursor.
OS   Capra hircus (Goat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Caprinae; Capra.
OX   NCBI_TaxID=9925;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Mammary gland;
RA   Paramasivam M., Saravanan K., Sharma S., Mohanty A.K., Yadav S.,
RA   Srinivasan A., Singh T.P.;
RT   "Goat mammary gland mRNA for a novel protein, BP40, complete coding
RT   region.";
RL   Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY081150; AAL87007.1; -.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   ProDom; PD000471; Glyco_hydro_18; 1.
DR   SMART; SM00636; Glyco_18; 1.
KW   Signal.
FT   SIGNAL        1     16       POTENTIAL.
SQ   SEQUENCE   383 AA;  42893 MW;  17655ED4BB4E9F5F CRC64;

  Query Match           100.0%;  Score 44;  DB 6;  Length 383;
  Best Local Similarity  100.0%;  Pred. No. 0.39;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FGRSFTLAS 9
              |||||||||
Db        261 FGRSFTLAS 269


RESULT 4
Q29411
ID   Q29411      PRELIMINARY;      PRT;   383 AA.
AC   Q29411;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   38 kDa heparin-binding glycoprotein.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Smooth muscle;
RX   MEDLINE=95286589; PubMed=7768902;
RA   Shackelton L.M., Mann D.M., Millis A.J.;
RT   "Identification of a 38-kDa heparin-binding glycoprotein (gp38k) in
RT   differentiating vascular smooth muscle cells as a member of a group of
RT   proteins associated with tissue remodeling.";
RL   J. Biol. Chem. 270:13076-13083(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Smooth muscle;
RA   Shackelton L.M., Mann D.M., Millis A.J.T.;
RL   Submitted (JAN-1995) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U19900; AAA86482.1; -.
DR   EMBL; Z47803; CAA87764.1; -.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   ProDom; PD000471; Glyco_hydro_18; 1.
DR   SMART; SM00636; Glyco_18; 1.
KW   Heparin-binding.
SQ   SEQUENCE   383 AA;  42443 MW;  3D1039F49910BDC4 CRC64;

  Query Match           100.0%;  Score 44;  DB 6;  Length 383;
  Best Local Similarity  100.0%;  Pred. No. 0.39;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FGRSFTLAS 9
              |||:|||||
Db        261 FGRSFTLAS 269


RESULT 5
O18949
ID   O18949      PRELIMINARY;      PRT;   332 AA.
AC   O18949;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Chitinase-like protein 1 (Fragment).
GN   CLP-1.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Recklies A.D., White C.;
RT   "Expression of chitinase-like protein 1 (CLP-1) in bovine
RT   chondrocytes.";
RL   Submitted (JUN-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF011373; AAB64304.1; -.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   ProDom; PD000471; Glyco_hydro_18; 1.
DR   SMART; SM00636; Glyco_18; 1.
FT   NON_TER       1       1
FT   NON_TER     332     332
SQ   SEQUENCE   332 AA;  37454 MW;  03F163323486408F CRC64;

  Query Match           93.2%;  Score 41;  DB 6;  Length 332;
  Best Local Similarity  88.9%;  Pred. No. 1.5;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FGRSFTLAS 9
              ||||:||||
Db        237 FGRSYTLAS 245
```

```
RESULT 6
Q9WTV1
ID   Q9WTV1      PRELIMINARY;      PRT;   352 AA.
AC   Q9WTV1;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Glycoprotein-39 (Fragment).
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Lewis;
RA   Wendling G., Boots A.M.H., van Eden W.;
RT   "Cloning of the rat homologue of Human Cartilage glycoprotein-39 a
RT   potential autoantigen in arthritis.";
RL   Submitted (APR-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF062038; AAD22610.1; -.
DR   HSSP; P07254; 1CTN.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   ProDom; PD000471; Glyco_hydro_18; 1.
DR   SMART; SM00636; Glyco_18; 1.
FT   NON_TER       1       1
FT   NON_TER     352     352
SQ   SEQUENCE   352 AA;  39391 MW;  CBDB991810AC9B6C CRC64;

  Query Match             93.2%;  Score 41;  DB 11;  Length 352;
  Best Local Similarity   88.9%;  Pred. No. 1.6;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FGRSFTLAS 9
              ||:||||||
Db        240 FGKSFTLAS 248


RESULT 7
Q99J84
ID   Q99J84      PRELIMINARY;      PRT;   381 AA.
AC   Q99J84;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Similar to chitinase 3-like 1 (Cartilage glycoprotein-39).
GN   CHI3L1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Strausberg R.;
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Strausberg R.;
RL   Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Strausberg R.;
RL   Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC005611; AAH05611.1; -.
DR   EMBL; BC003780; AAH03780.1; -.
DR   EMBL; BC004734; AAH04734.1; -.
DR   MGD; MGI:1340899; Chi3l1.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   ProDom; PD000471; Glyco_hydro_18; 1.
DR   SMART; SM00636; Glyco_18; 1.
SQ   SEQUENCE   381 AA;  42979 MW;  EF6588C5AE9D4450 CRC64;

  Query Match             93.2%;  Score 41;  DB 11;  Length 381;
  Best Local Similarity   88.9%;  Pred. No. 1.7;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FGRSFTLAS 9
              ||:||||||
Db        262 FGKSFTLAS 270


RESULT 8
Q9H3V8
ID   Q9H3V8      PRELIMINARY;      PRT;   387 AA.
AC   Q9H3V8;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Chitotriosidase precursor.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96064695; PubMed=7592832;
RA   Boot R.G., Renkema G.H., Strijland A., van Zonneveld A.J., Aerts J.M.;
RT   "Cloning of a cDNA encoding chitotriosidase, a human chitinase
RT   produced by macrophages.";
RL   J. Biol. Chem. 270:26252-26256(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98421482; PubMed=9748235;
RA   Boot R.G., Renkema G.H., Verhoek M., Strijland A., Bliek J.,
RA   de Meulemeester T.M., Mannens M.M., Aerts J.M.;
RT   "The human chitotriosidase gene. Nature of inherited enzyme
RT   deficiency.";
RL   J. Biol. Chem. 273:25680-25685(1998).
DR   EMBL; U62662; AAG10644.1; -.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   InterPro; IPR001579; Glyco_hydro_18/2.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   ProDom; PD000471; Glyco_hydro_18; 1.
DR   SMART; SM00636; Glyco_18; 1.
DR   PROSITE; PS01095; CHITINASE_18; 1.
KW   Signal.
FT   SIGNAL        1     21       POTENTIAL.
FT   CHAIN        22    387       CHITOTRIOSIDASE.
SQ   SEQUENCE   387 AA;  43133 MW;  03A272B8BC5ROD71 CRC64;

  Query Match             93.2%;  Score 41;  DB 4;  Length 387;
  Best Local Similarity   88.9%;  Pred. No. 1.7;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FGRSFTLAS 9
              :||||||||
Db        267 YGRSFTLAS 275


RESULT 9
Q8BKL8
ID   Q8BKL8      PRELIMINARY;      PRT;   389 AA.
AC   Q8BKL8;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Similar to chitinase 3-like 1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Dorsal root ganglion;
RX   MEDLINE=22354683; PubMed=12466851;
```

```
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
DR   EMBL; AK051475; BAC34654.1; -.
SQ   SEQUENCE   389 AA;  43893 MW;  9E7D66069B233834 CRC64;

  Query Match             93.2%;  Score 41;  DB 11;  Length 389;
  Best Local Similarity   88.9%;  Pred. No. 1.8;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FGRSFTLAS 9
              |.:||||||
Db        270 FGKSFTLAS 278


RESULT 10
Q9D7Q1
ID   Q9D7Q1      PRELIMINARY;      PRT;   396 AA.
AC   Q9D7Q1;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   2300002L19Rik protein.
GN   2300002L19RIK.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Tongue;
RX   MEDLINE=21065660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
DR   EMBL; AK009012; BAB25025.1; -.
DR   MGD; MGI:1919134; 2300002L19Rik.
DR   InterPro; IPR000677; 2S_Globulin.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   InterPro; IPR001579; Glyco_hydro_18/2.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   PRINTS; PR00551; 2SGLOBULIN.
DR   ProDom; PD000471; Glyco_hydro_18; 1.
DR   SMART; SM00636; Glyco_18; 1.
DR   PROSITE; PS01095; CHITINASE_18; 1.
SQ   SEQUENCE   396 AA;  43413 MW;  32BB674B61AF7123 CRC64;

  Query Match             93.2%;  Score 41;  DB 11;  Length 396;
  Best Local Similarity   88.9%;  Pred. No. 1.8;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FGRSFTLAS 9
              |||||||||
Db        267 YGRSFTLAS 275


RESULT 11
Q13231
ID   Q13231      PRELIMINARY;      PRT;   466 AA.
AC   Q13231;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Chitotriosidase precursor.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96064695; PubMed=7592832;
RA   Boot R.G., Renkema G.H., Strijland A., van Zonneveld A.J., Aerts J.M.;
RT   "Cloning of a cDNA encoding chitotriosidase, a human chitinase
RT   produced by macrophages.";
RL   J. Biol. Chem. 270:26252-26256(1995).
DR   EMBL; U29615; AAC50246.1; -.
DR   Genew; HGNC:1936; CHIT1.
DR   InterPro; IPR002557; Chitin_bind_PerA.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   InterPro; IPR001579; Glyco_hydro_18/2.
DR   Pfam; PF01607; CBM_14; 1.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   ProDom; PD000471; Glyco_hydro_18; 1.
DR   SMART; SM00494; ChtBD2; 1.
DR   SMART; SM00636; Glyco_18; 1.
DR   PROSITE; PS01095; CHITINASE_18; 1.
KW   Signal.
FT   SIGNAL        1     21       POTENTIAL.
FT   CHAIN        22    466       CHITOTRIOSIDASE.
SQ   SEQUENCE   466 AA;  51681 MW;  B4312D1E885E3B6D CRC64;

  Query Match             93.2%;  Score 41;  DB 4;  Length 466;
  Best Local Similarity   88.9%;  Pred. No. 2.1;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FGRSFTLAS 9
              :|||||||||
Db        267 YGRSFTLAS 275


RESULT 12
Q8MSI4
ID   Q8MSI4      PRELIMINARY;      PRT;   268 AA.
AC   Q8MSI4;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   GH13872p.
GN   CG9357.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RA   Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA   Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA   George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA   Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA   Patel S., Phouanenavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA   Celniker S.;
RL   Submitted (JUN-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY118784; AAM50644.1; -.
DR   FlyBase; FBgn0034580; CG9357.
DR   InterPro; IPR002557; Chitin_bind_PerA.
```

```
DR   InterPro; IPR001179; FKBP_PPIase.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF01607; CBM_14; 1.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   ProDom; PD000471; Glyco_hydro_18; 1.
DR   SMART; SM00494; ChtBD2; 1.
DR   SMART; SM00636; Glyco_18; 1.
DR   PROSITE; PS00453; FKBP_PPIASE_1; 1.
SQ   SEQUENCE   268 AA;  29678 MW;  5DEBDC8903CB2D1E CRC64;

  Query Match             86.4%;  Score 38;  DB 5;  Length 268;
  Best Local Similarity   77.8%;  Pred. No. 5.3;
  Matches    7;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FGRSFTLAS 9
              :|||||||:
Db         72 YGRSFTLAT 80


RESULT 13
Q8ISH5
ID   Q8ISH5      PRELIMINARY;      PRT;   431 AA.
AC   Q8ISH5;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Chitinase.
OS   Araneus ventricosus.
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;
OC   Araneomorphae; Entelegynae; Araneoidea; Araneidae; Araneus.
OX   NCBI_TaxID=182803;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Han J.H., Kim S.R., Sohn H.D., Jin B.R.;
RT   "Molecular cloning of a cDNA encoding the chitinase from the spider,
RT   Araneus ventricosus.";
RL   Submitted (JUN-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY120879; AAN39100.1; -.
SQ   SEQUENCE   431 AA;  47238 MW;  929439397B9BC923 CRC64;

  Query Match             86.4%;  Score 38;  DB 5;  Length 431;
  Best Local Similarity   77.8%;  Pred. No. 8.7;
  Matches    7;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FGRSFTLAS 9
              :|||:|||:
Db        274 YGRSFTLAN 282


RESULT 14
Q9W2M7
ID   Q9W2M7      PRELIMINARY;      PRT;   460 AA.
AC   Q9W2M7;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   CG9357 protein.
GN   CG9357.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
DR   EMBL; AE003452; AAF46663.1; -.
DR   HSSP; P07254; 1CTN.
DR   FlyBase; FBgn0034580; CG9357.
DR   InterPro; IPR002557; Chitin_bind_PerA.
DR   InterPro; IPR001179; FKBP_PPIase.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   InterPro; IPR001579; Glyco_hydro_18/2.
DR   Pfam; PF01607; CBM_14; 1.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   ProDom; PD000471; Glyco_hydro_18; 1.
DR   SMART; SM00494; ChtBD2; 1.
DR   SMART; SM00636; Glyco_18; 1.
DR   PROSITE; PS01095; CHITINASE_18; 1.
DR   PROSITE; PS00453; FKBP_PPIASE_1; 1.
KW   Glycosidase; Hydrolase.
SQ   SEQUENCE   460 AA;  51055 MW;  8EF4F126F1DD8D9B CRC64;

  Query Match             86.4%;  Score 38;  DB 5;  Length 460;
  Best Local Similarity   77.8%;  Pred. No. 9.3;
  Matches    7;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FGRSFTLAS 9
              :|||:|||:
Db        264 YGRSFTLAT 272


RESULT 15
O44079
ID   O44079      PRELIMINARY;      PRT;   525 AA.
AC   O44079;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Chitinase.
GN   AGCH1-1.
OS   Anopheles gambiae (African malaria mosquito).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX   NCBI_TaxID=7165;
RN   [1]
RP   SEQUENCE FROM N.A.
```

```
RC   TISSUE=Gut;
RX   MEDLINE=98030563; PubMed=9360958;
RA   Shen Z., Jacobs-Lorena M.;
RT   "Characterization of a novel gut-specific chitinase gene from the
RT   human malaria vector Anopheles gambiae.";
RL   J. Biol. Chem. 272:28895-28900(1997).
DR   EMBL; AF008575; AAB87764.1; -.
DR   InterPro; IPR002557; Chitin_bind_PerA.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   InterPro; IPR001579; Glyco_hydro_18/2.
DR   Pfam; PF01607; CBM_14; 1.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   ProDom; PD000471; Glyco_hydro_18; 1.
DR   SMART; SM00494; ChtBD2; 1.
DR   SMART; SM00636; Glyco_18; 1.
DR   PROSITE; PS01095; CHITINASE_18; 1.
KW   Glycosidase; Hydrolase.
SQ   SEQUENCE   525 AA;  57211 MW;  3234360EEFF36165 CRC64;

  Query Match             66.4%;  Score 38;  DB 5;  Length 525;
  Best Local Similarity   77.8%;  Pred. No. 11;
  Matches    7;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FQRSFTLAS 9
              :|:|:|||:|
Db        279 YGRNFTLAS 287


Search completed: September 24, 2003, 13:47:21
Job time : 31.2247 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2003, 13:25:57 ; Search time 36.1011 Seconds
(without alignments)
39.570 Million cell updates/sec

Title: US-09-744-282-2
Perfect score: 43
Sequence: 1 FTLASSETG 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	43	100.0	9	17	AAR95072	Human cartilage gl
2	43	100.0	9	21	AAY69227	Peptide fragment o
3	43	100.0	9	23	AAE20213	Human cartilage gl
4	43	100.0	11	23	AAE20206	Human cartilage gl
5	43	100.0	11	23	AAE20210	Human cartilage gl
6	43	100.0	11	23	AAE20212	Human cartilage gl
7	43	100.0	12	23	AAE20208	Human cartilage gl
8	43	100.0	13	17	AAR95076	Peptide contg. hC
9	43	100.0	13	19	AAW61597	Human Cartilage gp
10	43	100.0	13	21	AAY69231	Peptide fragment o
11	43	100.0	13	21	AAY77401	Human cartilage gl
12	43	100.0	13	22	AAU03045	Human cartilage gp
13	43	100.0	13	22	AAU03046	Human cartilage gp
14	43	100.0	13	22	AAU03047	Human cartilage gp
15	43	100.0	13	22	AAU03048	Human cartilage gp
16	43	100.0	13	22	AAU03054	Human cartilage gp
17	43	100.0	13	22	AAU03055	Human cartilage gp
18	43	100.0	13	22	AAU03057	Human cartilage gp
19	43	100.0	13	22	AAU03058	Human cartilage gp
20	43	100.0	13	22	AAU03059	Human cartilage gp
21	43	100.0	13	22	AAU03062	Human cartilage gp
22	43	100.0	13	22	AAU03064	Human cartilage gp
23	43	100.0	13	22	AAU03065	Human cartilage gp
24	43	100.0	13	22	AAU03066	Human cartilage gp
25	43	100.0	13	22	AAU03067	Human cartilage gp
26	43	100.0	13	22	AAU03072	Human cartilage gp
27	43	100.0	13	22	AAU03073	Human cartilage gp
28	43	100.0	13	23	AAE20207	Human cartilage gl
29	43	100.0	13	23	AAE20214	Human cartilage gl
30	43	100.0	13	23	AAE20215	Human cartilage gl
31	43	100.0	13	23	AAE20216	Human cartilage gl
32	43	100.0	13	23	AAE20217	Human cartilage gl
33	43	100.0	13	23	AAE20218	Human cartilage gl
34	43	100.0	13	23	AAE20220	Human cartilage gl
35	43	100.0	14	22	AAU03070	Human cartilage gp
36	43	100.0	14	22	AAU03074	Human cartilage gp
37	43	100.0	15	19	AAW61598	Human Cartilage gp
38	43	100.0	16	18	AAW45140	Human cartilage gl
39	43	100.0	16	19	AAW61599	Human Cartilage gp
40	43	100.0	105	22	AAU02074	Synthetic human ta
41	43	100.0	199	22	ABB56452	Human single chain
42	43	100.0	199	22	ABB56456	Human single chain
43	43	100.0	208	22	ABB56475	Human single chain
44	43	100.0	209	22	ABB56454	Human single chain
45	43	100.0	210	22	ABB56442	Human single chain

ALIGNMENTS

RESULT 1
AAR95072
ID AAR95072 standard; peptide; 9 AA.
XX
AC AAR95072;
XX
DT 17-DEC-1996 (first entry)
XX
DE Human cartilage glycoprotein 39 autoantigen derived peptide (II).
XX
KW Human; cartilage; glycoprotein 39; hC gp-39; autoantigen;
KW induction; T-cell; tolerance; mediated; destruction;
KW rheumatoid arthritis; disease model; diagnosis.
XX
OS Homo sapiens.
XX
PN WO9613517-A1.
XX
PD 09-MAY-1996.
XX
PF 25-OCT-1995; 95WO-EP04201.
XX
PR 07-APR-1995; 95EP-0200886.
PR 27-OCT-1994; 94EP-0203128.
XX
PA (AKZO) AKZO NOBEL NV.
XX
PI Boots AMH, Verheijden GFM;
XX
DR WPI; 1996-239447/24.
XX
PT Peptide(s) derived from human cartilage gp-39 autoantigen - used

One or more of your sequences had more than forty-five 100% matches when searched against requested databases. More scores and alignments can be saved, or sequences retrieved can be limited by length. If either of these options would be helpful, please call me.

Barb O'Bryen 308-4291

PT to induce specific T cell tolerance, esp. for treatment of
PT rheumatoid arthritis, also useful in producing arthritic animal
PT models
XX
PS Claim 1; Page 28; 37pp; English.
XX
CC The present peptide, derived from the human cartilage glycoprotein
CC 39 (hC gp-39) autoantigen, can be used to induce specific T-cell
CC tolerance to hC gp-39 in a patient with T-cell mediated cartilage
CC destruction, esp. rheumatoid arthritis (RA). It may also be used to
CC induce RA in mice, etc. (human disease models), and to diagnose the
CC presence of activated T-cells autoreactive with hC gp-39, i.e. to
CC identify patients suitable for tolerising treatment.
CC Therapeutically the peptide is given in a dose of 0.1-100 microg/kg
CC by injection, or 10-50 microg/kg to produce a disease model. The
CC peptide has a specific tolerising effect on autoreactive T-cells,
CC but leaves other components of the immune system intact, and
CC therefore should have few side effects. In an assay to determine
CC the peptide's binding to HLA-DR mols., specifically DR4Dw4 (known
CC to be associated with an increased risk of RA development), the
CC peptide gave a DOT blot IC(50) (micromolar) value of 0.08 compared
CC to 0.55 for the control peptide IHA(307-319)F.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 43; DB 17; Length 9;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FTLASSETG 9
|||||||||
Db 1 FTLASSETG 9

RESULT 2
AAY69227
ID AAY69227 standard; peptide; 9 AA.
XX
AC AAY69227;
XX
DT 30-MAY-2000 (first entry)
XX
DE Peptide fragment of human cartilage gp-39 (HC gp-39) protein.
XX
KW Human cartilage gp-39 protein; HC gp-39 protein; T-cell activator;
KW inflammatory disease; lymphocyte; antigen; bystander suppression;
KW antigen-specific T-cell tolerance; autoimmune disorder;
KW inflammatory disease; rheumatoid arthritis;
KW autoimmune inflammatory disease; autoantigen.
XX
OS Homo sapiens.
XX
PN WO200004917-A2.
XX
PD 03-FEB-2000.
XX
PF 19-JUL-1999; 99WO-EP05331.
XX
PR 23-JUL-1998; 98EP-0202471.
XX
PA (ALKU) AKZO NOBEL NV.
XX
PI Miltenburg AMM, Boots AMH;
XX
DR WPI; 2000-182537/16.
XX
PT Use of cartilage protein or peptides for treating inflammatory diseases
PT especially rheumatoid arthritis -
XX
PS Claim 2; Page 17; 29pp; English.
XX
CC AAY69226-33 represent fragments of the human cartilage gp-39 (HC gp-39)
CC protein. The peptides are T-cell activators. The peptides are used for
CC preventing inflammatory diseases, by modulating lymphocytes reactive
CC to antigens other than HC gp-39, present in the same tissue of
CC HC gp-39. By the induction of antigen-specific T-cell tolerance,
CC autoimmune disorders can be treated by bystander suppression. The
CC HC gp-39 peptides are useful for preventing inflammatory diseases
CC especially rheumatoid arthritis. The fragments may also be used for
CC treating autoimmune inflammatory diseases. The HC gp-39 peptides are
CC especially beneficial in autoimmune conditions in which the
CC autoantigens associated with the condition are unknown.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 43; DB 21; Length 9;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FTLASSETG 9
|||||||||
Db 1 FTLASSETG 9

RESULT 3
AAE20213
ID AAE20213 standard; peptide; 9 AA.
XX
AC AAE20213;
XX
DT 18-JUN-2002 (first entry)
XX
DE Human cartilage glycoprotein (HC gp)-39 peptide (residues 265-273).
XX
KW Human cartilage glycoprotein-39; HC gp-39; autoimmune disorder; MHC;
KW major histocompatibility complex; insulin-dependent diabetes mellitus;
KW multiple sclerosis; psoriasis; myasthenia gravis; rheumatoid arthritis;
KW immunosuppressive; neuroprotective; therapy.
XX
OS Homo sapiens.
XX
PN WO200214870-A2.
XX
PD 21-FEB-2002.
XX
PF 08-AUG-2001; 2001WO EP09136.
XX
PR 14-AUG-2000; 2000EP-0202844.
XX
PA (ALKU) AKZO NOBEL NV.
XX
PI Steenbakkers PGA;
XX
DR WPI; 2002-269214/31.
XX
PT Method for diagnosing an autoimmune disease activity by detecting the
PT presence of an autoimmune specific major histocompatibility complex
PI (MHC)-peptide complex in a patient using antibodies specific against
PT MHC-peptide complexes -
XX
PS Example 4; Page 23; 43pp; English.
XX
CC The patent discloses a method for diagnosing an autoimmune disease
CC activity. The method involves detecting the presence of an autoimmune
CC specific major histocompatibility complex (MHC)-peptide complex in a
CC patient suffering from an autoimmune disease with antibodies or its
CC antigen binding domains that specifically bind to a MHC-gp39-derived
CC peptide. The method and antibodies are useful for diagnosing auto-
CC immune diseases or purifying autoimmune specific MHC-peptide complexes.
CC They are also useful in therapy or for the manufacture of pharmaceutical
CC preparation for the treatment of autoimmune disorders such as insulin-
CC dependent diabetes mellitus, multiple sclerosis, myasthenia gravis,
CC psoriasis or rheumatoid arthritis. The present sequence is human
CC cartilage glycoprotein (HC gp)-39 peptide (residues 265-273). This
CC peptide is used in the exemplification of the invention.
XX

```
SQ   Sequence   9 AA;

  Query Match            100.0%;  Score 43;  DB 23;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 9.3e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTLASSETG 9
              |||||||||
Db          1 FTLASSETG 9


RESULT 4
AAE20205
ID   AAE20205 standard; peptide; 11 AA.
XX
AC   AAE20205;
XX
DT   18-JUN-2002  (first entry)
XX
DE   Human cartilage glycoprotein (HC gp)-39 peptide (residues 263-273).
XX
KW   Human cartilage glycoprotein-39; HC gp-39; autoimmune disorder; MHC;
KW   major histocompatibility complex; insulin-dependent diabetes mellitus;
KW   multiple sclerosis; psoriasis; myasthenia gravis; rheumatoid arthritis;
KW   immunosuppressive; neuroprotective; therapy.
XX
OS   Homo sapiens.
XX
PN   WO200214870-A2.
XX
PD   21-FEB-2002.
XX
PF   08-AUG-2001; 2001WO-EP09136.
XX
PR   14-AUG-2000; 2000EP-0202844.
XX
PA   (ALKU ) AKZO NOBEL NV.
XX
PI   Steenbakkers PGA;
XX
DR   WPI; 2002-269214/31.
XX
PT   Method for diagnosing an autoimmune disease activity by detecting the
PT   presence of an autoimmune specific major histocompatibility complex
PT   (MHC)-peptide complex in a patient using antibodies specific against
PT   MHC-peptide complexes -
XX
PS   Example 4; Page 23; 43pp; English.
XX
CC   The patent discloses a method for diagnosing an autoimmune disease
CC   activity. The method involves detecting the presence of an autoimmune
CC   specific major histocompatibility complex (MHC)-peptide complex in a
CC   patient suffering from an autoimmune disease with antibodies or its
CC   antigen binding domains that specifically bind to a MHC-gp39-derived
CC   peptide. The method and antibodies are useful for diagnosing auto-
CC   immune diseases or purifying autoimmune specific MHC-peptide complexes.
CC   They are also useful in therapy or for the manufacture of pharmaceutical
CC   preparation for the treatment of autoimmune disorders such as insulin-
CC   dependent diabetes mellitus, multiple sclerosis, myasthenia gravis,
CC   psoriasis or rheumatoid arthritis. The present sequence is human
CC   cartilage glycoprotein (HC gp)-39 peptide (residues 263-273). This
CC   peptide is used in the exemplification of the invention.
XX
SQ   Sequence   11 AA;

  Query Match            100.0%;  Score 43;  DB 23;  Length 11;
  Best Local Similarity  100.0%;  Pred. No. 0.0068;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTLASSETG 9
              |||||||||
Db          3 FTLASSETG 11


RESULT 5
AAE20210
ID   AAE20210 standard; peptide; 11 AA.
XX
AC   AAE20210;
XX
DT   18-JUN-2002  (first entry)
XX
DE   Human cartilage glycoprotein (HC gp)-39 peptide (residues 265-275).
XX
KW   Human cartilage glycoprotein-39; HC gp-39; autoimmune disorder; MHC;
KW   major histocompatibility complex; insulin-dependent diabetes mellitus;
KW   multiple sclerosis; psoriasis; myasthenia gravis; rheumatoid arthritis;
KW   immunosuppressive; neuroprotective; therapy.
XX
OS   Homo sapiens.
XX
PN   WO200214870-A2.
XX
PD   21-FEB-2002.
XX
PF   08-AUG-2001; 2001WO-EP09136.
XX
PR   14-AUG-2000; 2000EP-0202844.
XX
PA   (ALKU ) AKZO NOBEL NV.
XX
PI   Steenbakkers PGA;
XX
DR   WPI; 2002-269214/31.
XX
PT   Method for diagnosing an autoimmune disease activity by detecting the
PT   presence of an autoimmune specific major histocompatibility complex
PT   (MHC)-peptide complex in a patient using antibodies specific against
PT   MHC-peptide complexes -
XX
PS   Example 4; Page 23; 43pp; English.
XX
CC   The patent discloses a method for diagnosing an autoimmune disease
CC   activity. The method involves detecting the presence of an autoimmune
CC   specific major histocompatibility complex (MHC)-peptide complex in a
CC   patient suffering from an autoimmune disease with antibodies or its
CC   antigen binding domains that specifically bind to a MHC-gp39-derived
CC   peptide. The method and antibodies are useful for diagnosing auto-
CC   immune diseases or purifying autoimmune specific MHC-peptide complexes.
CC   They are also useful in therapy or for the manufacture of pharmaceutical
CC   preparation for the treatment of autoimmune disorders such as insulin-
CC   dependent diabetes mellitus, multiple sclerosis, myasthenia gravis,
CC   psoriasis or rheumatoid arthritis. The present sequence is human
CC   cartilage glycoprotein (HC gp)-39 peptide (residues 265-275). This
CC   peptide is used in the exemplification of the invention.
XX
SQ   Sequence   11 AA;

  Query Match            100.0%;  Score 43;  DB 23;  Length 11;
  Best Local Similarity  100.0%;  Pred. No. 0.0068;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTLASSETG 9
              |||||||||
Db          1 FTLASSETG 9


RESULT 6
AAE20212
ID   AAE20212 standard; peptide; 11 AA.
XX
AC   AAE20212;
XX
DT   18-JUN-2002  (first entry)
XX
DE   Human cartilage glycoprotein (HC gp)-39 peptide (residues 264-274).
```

```
XX
KW   Human cartilage glycoprotein-39; HC gp-39; autoimmune disorder; MHC;
KW   major histocompatibility complex; insulin-dependent diabetes mellitus;
KW   multiple sclerosis; psoriasis; myasthenia gravis; rheumatoid arthritis;
KW   immunosuppressive; neuroprotective; therapy.
XX
OS   Homo sapiens.
XX
PN   WO200214870-A2.
XX
PD   21-FEB-2002.
XX
PF   08-AUG-2001; 2001WO-EP09136.
XX
PR   14-AUG-2000; 2000EP-0202844.
XX
PA   (ALKU ) AKZO NOBEL NV.
XX
PI   Steenbakkers PGA;
XX
DR   WPI; 2002-269214/31.
XX
PT   Method for diagnosing an autoimmune disease activity by detecting the
PT   presence of an autoimmune specific major histocompatibility complex
PT   (MHC)-peptide complex in a patient using antibodies specific against
PT   MHC-peptide complexes -
XX
PS   Example 4; Page 23; 43pp; English.
XX
CC   The patent discloses a method for diagnosing an autoimmune disease
CC   activity. The method involves detecting the presence of an autoimmune
CC   specific major histocompatibility complex (MHC)-peptide complex in a
CC   patient suffering from an autoimmune disease with antibodies or its
CC   antigen binding domains that specifically bind to a MHC-gp39-derived
CC   peptide. The method and antibodies are useful for diagnosing auto-
CC   immune diseases or purifying autoimmune specific MHC-peptide complexes.
CC   They are also useful in therapy or for the manufacture of pharmaceutical
CC   preparation for the treatment of autoimmune disorders such as insulin-
CC   dependent diabetes mellitus, multiple sclerosis, myasthenia gravis,
CC   psoriasis or rheumatoid arthritis. The present sequence is human
CC   cartilage glycoprotein (HC gp)-39 peptide (residues 264-274). This
CC   peptide is used in the exemplification of the invention.
XX
SQ   Sequence   11 AA;

  Query Match             100.0%;  Score 43;  DB 23;  Length 11;
  Best Local Similarity   100.0%;  Pred. No. 0.0068;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTLASSETG 9
              |||||||||
Db          2 FTLASSETG 10


RESULT 7
AAE20208
ID   AAE20208 standard; peptide; 12 AA.
XX
AC   AAE20208;
XX
DT   18-JUN-2002  (first entry)
XX
DE   Human cartilage glycoprotein (HC gp)-39 peptide (residues 263-274).
XX
KW   Human cartilage glycoprotein-39; HC gp-39; autoimmune disorder; MHC;
KW   major histocompatibility complex; insulin-dependent diabetes mellitus;
KW   multiple sclerosis; psoriasis; myasthenia gravis; rheumatoid arthritis;
KW   immunosuppressive; neuroprotective; therapy.
XX
OS   Homo sapiens.
XX
PN   WO200214870-A2.
XX
PD   21-FEB-2002.
XX
PF   08-AUG-2001; 2001WO-EP09136.
XX
PR   14-AUG-2000; 2000EP-0202844.
XX
PA   (ALKU ) AKZO NOBEL NV.
XX
PI   Steenbakkers PGA;
XX
DR   WPI; 2002-269214/31.
XX
PT   Method for diagnosing an autoimmune disease activity by detecting the
PT   presence of an autoimmune specific major histocompatibility complex
PT   (MHC)-peptide complex in a patient using antibodies specific against
PT   MHC-peptide complexes -
XX
PS   Example 4; Page 23; 43pp; English.
XX
CC   The patent discloses a method for diagnosing an autoimmune disease
CC   activity. The method involves detecting the presence of an autoimmune
CC   specific major histocompatibility complex (MHC)-peptide complex in a
CC   patient suffering from an autoimmune disease with antibodies or its
CC   antigen binding domains that specifically bind to a MHC-gp39-derived
CC   peptide. The method and antibodies are useful for diagnosing auto-
CC   immune diseases or purifying autoimmune specific MHC-peptide complexes.
CC   They are also useful in therapy or for the manufacture of pharmaceutical
CC   preparation for the treatment of autoimmune disorders such as insulin-
CC   dependent diabetes mellitus, multiple sclerosis, myasthenia gravis,
CC   psoriasis or rheumatoid arthritis. The present sequence is human
CC   cartilage glycoprotein (HC gp)-39 peptide (residues 263-274). This
CC   peptide is used in the exemplification of the invention.
XX
SQ   Sequence   12 AA;

  Query Match             100.0%;  Score 43;  DB 23;  Length 12;
  Best Local Similarity   100.0%;  Pred. No. 0.0076;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTLASSETG 9
              |||||||||
Db          3 FTLASSETG 11


RESULT 8
AAR95076
ID   AAR95076 standard; peptide; 13 AA.
XX
AC   AAR95076;
XX
DT   17-DEC-1996  (first entry)
XX
DE   Peptide contg. hC gp-39 autoantigen derived peptide (II).
XX
KW   Human; cartilage; glycoprotein 39; hC gp-39; autoantigen;
KW   induction; T-cell; tolerance; mediated; destruction;
KW   rheumatoid arthritis; disease model; diagnosis.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         3..11
FT                   /note= "HC gp-39 peptide (II)"
XX
PN   WO9613517-A1.
XX
PD   09-MAY-1996.
XX
PF   25-OCT-1995;   95WO-EP04201.
XX
PR   07-APR-1995;   95EP-0200886.
PR   27-OCT-1994;   94EP-0203128.
XX
```

```
PA   (ALKU ) AKZO NOBEL NV.
XX
PI   Boots AMH,  Verheijden GFM;
XX
DR   WPI; 1996-239447/24.
XX
PT   Peptide(s) derived from human cartilage gp-39 autoantigen - used
PT   to induce specific T cell tolerance, esp. for treatment of
PT   rheumatoid arthritis, also useful in producing arthritic animal
PT   models
XX
PS   Claim 2; Page 28; 37pp; English.
XX
CC   The present peptide, contg. a peptide derived from the human
CC   cartilage glycoprotein 39 (hC gp-39) autoantigen, can be used to
CC   induce specific T-cell tolerance to hC gp-39 in a patient with
CC   T-cell mediated cartilage destruction, esp. rheumatoid arthritis
CC   (RA). It may also be used to induce RA in mice, etc. (human disease
CC   models), and to diagnose the presence of activated T-cells
CC   autoreactive with hC gp-39, i.e. to identify patients suitable for
CC   tolerising treatment.
CC   Therapeutically the peptide is given in a dose of 0.1-100 microg/kg
CC   by injection, or 10-50 microg/kg to produce a disease model. The
CC   peptide has a specific tolerising effect on autoreactive T-cells,
CC   but leaves other components of the immune system intact, and
CC   therefore should have few side effects.
XX
SQ   Sequence   13 AA;

  Query Match             100.0%;  Score 43;  DB 17;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 0.0083;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTLASSETG 9
              |||||||||
Db          3 FTLASSETG 11


RESULT 9
AAW61597
ID   AAW61597 standard; protein; 13 AA.
XX
AC   AAW61597;
XX
DT   27-OCT-1998  (first entry)
XX
DE   Human Cartilage gp-39 epitope (1).
XX
KW   Human; cartilage gp-39; monoclonal antibody; cell surface antigen;
KW   rheumatoid arthritis; T-cell receptor.
XX
OS   Homo sapiens.
XX
PN   EP856520-A1.
XX
PD   05-AUG-1998.
XX
PF   02-DEC-1997;   97EP-0203769.
XX
PR   27-JUN-1997;   97EP-0201972.
PR   06-DEC-1996;   96EP-0203465.
XX
PA   (ALKU ) AKZO NOBEL NV.
PA   (ALKU ) AKZO NV.
XX
PI   Steenbakkers PGA;
XX
DR   WPI; 1998-400866/35.
XX
PT   Preparation of monoclonal antibodies against cell surface antigens
PT   used as diagnostic agents and to treat rheumatoid arthritis -
PT   comprises immunising, then enriching, enriching and cloning B cells
PI   and then selecting and cloning hybridoma
XX
PS   Examples; Page 5; 23pp; English.
XX
CC   The human cartilage gp-39 peptides AAW61597-W61599 were used in the
CC   preparation of a monoclonal antibody against a cell surface antigen.
CC   The antibody can be used for treatment of rheumatoid arthritis.  It
CC   is also useful against cell surface antigens and antibodies reactive
CC   with the clonotypic structure of T-cell receptors are used as diagnostic
CC   reagents.  The process allows antibodies to be raised against
CC   antigens which are present in very low amounts or which are
CC   conformationally unstable without the need to screen very large numbers
CC   of hybridomas and without requiring high purity Antigen for enrichment.
XX
SQ   Sequence   13 AA;

  Query Match             100.0%;  Score 43;  DB 19;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 0.0083;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTLASSETG 9
              |||||||||
Db          3 FTLASSETG 11


RESULT 10
AAY69231
ID   AAY69231 standard; peptide; 13 AA.
XX
AC   AAY69231;
XX
DT   30-MAY-2000  (first entry)
XX
DE   Peptide fragment of human cartilage gp-39 (HC gp-39) protein.
XX
KW   Human cartilage gp-39 protein; HC gp-39 protein; T-cell activator;
KW   inflammatory disease; lymphocyte; antigen; bystander suppression;
KW   antigen-specific T-cell tolerance; autoimmune disorder;
KW   inflammatory disease; rheumatoid arthritis;
KW   autoimmune inflammatory disease; autoantigen.
XX
OS   Homo sapiens.
XX
PN   WO200004917-A2.
XX
PD   03-FEB-2000.
XX
PF   19-JUL-1999;   99WO-EP05331.
XX
PR   23-JUL-1998;   98EP-0202471.
XX
PA   (ALKU ) AKZO NOBEL NV.
XX
PI   Miltenburg AMM,  Boots AMH;
XX
DR   WPI; 2000-182537/16.
XX
PT   Use of cartilage protein or peptides for treating inflammatory diseases
PT   especially rheumatoid arthritis -
XX
PS   Claim 2; Page 17; 29pp; English.
XX
CC   AAY69226-33 represent fragments of the human cartilage gp-39 (HC gp-39)
CC   protein. The peptides are T-cell activators. The peptides are used for
CC   preventing inflammatory diseases, by modulating lymphocytes reactive
CC   to antigens other than HC gp-39, present in the same tissue of
CC   HC gp-39. By the induction of antigen-specific T-cell tolerance,
CC   autoimmune disorders can be treated by bystander suppression. The
CC   HC gp-39 peptides are useful for preventing inflammatory diseases
CC   especially rheumatoid arthritis. The fragments may also be used for
CC   treating autoimmune inflammatory diseases. The HC gp-39 peptides are
CC   especially beneficial in autoimmune conditions in which the
CC   autoantigens associated with the condition are unknown.
XX
```

```
SQ   Sequence   13 AA;

  Query Match             100.0%;  Score 43;  DB 21;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 0.0083;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTLASSETG 9
              |||||||||
Db          3 FTLASSETG 11


RESULT 11
AAY77401
ID   AAY77401 standard; peptide; 13 AA.
XX
AC   AAY77401;
XX
DT   22-MAY-2000  (first entry)
XX
DE   Human cartilage glycoprotein-39 (HC gp-39), residues 263-275.
XX
KW   Human cartilage glycoprotein-39; HC gp-39; YKL-40; epitope;
KW   autoreactive T-cell; tolerisation; desensitisation; autoantigen;
KW   rheumatoid arthritis; human chondrocyte protein; YKL-39;
KW   bystander suppression; autoimmune disorder; Graves' disease;
KW   primary glomerulonephritis; inflammatory bowel disease; diagnosis.
XX
OS   Homo sapiens.
XX
PN   WO200005254-A2.
XX
PD   03-FEB-2000.
XX
PF   16-JUL-1999;   99WO-EP05050.
XX
PR   23-JUL-1998;   98EP-0202470.
XX
PA   (ALKU ) AKZO NOBEL NV.
XX
PI   Verheijden GFM,  Boots AMH;
XX
DR   WPI; 2000-182644/16.
XX
PT   New T-cell epitope peptides useful for inducing specific T-cell
PT   tolerance to an autoantigen for treating autoimmune disorders,
PT   especially rheumatoid arthritis  -
XX
PS   Example 1; Page 10; 26pp; English.
XX
CC   The inevntion relates to peptides constituting amino acids 268-276 and
CC   amino acids 266-278 of human chondrocyte protein YKL-39
CC   (AAY77399-Y77400, respectively). The peptides of the invention may be
CC   used in pharmaceutical compositions for the induction of specific T-cell
CC   tolerance (desensitisation) to an autoantigen in patients
CC   suffering from autoimmune disorders, especially rheumatoid arthritis. A
CC   frequently found autoantigen in rheumatoid arthritis sufferers is human
CC   cartilage glycoprotein-39 (HC gp-39, also known as YKL-40). Several
CC   epitopes of HC gp-39 are recognised by peripheral blood T-cells from
CC   rheumatoid arthritis patients; one such epitope, HC gp-39 residues
CC   263-275, was recognised by T-cells from 8 out of 18 patients tested.
CC   YKL-39 shares significant sequence identity to hc gp-39. YKL-39 residues
CC   266-278 (AAY77400) constitutes a mimicry epitope of HC gp-39 (263-275).
CC   and may be used for desensitisation of T-cells with reactivity to HC
CC   gp-39 (263-275) in rheumatoid arthritis patients. Rheumatoid arthritis
CC   may be treated by the systemic administration of compositions comprising
CC   the peptides of the invention, optionally in combination with existing
CC   treatments. The peptides may also be used for modulating lymphocytes
CC   that are reactive to autoantigens other than YKL-39 and HC gp-39 but
CC   which are present in the same tissue as these autoantigens. By the
CC   induction of antigen-specific T-cell tolerance, autoimmune disorders can
CC   be treated by bystander suppression. The conditions that may be treated
CC   via this approach include Graves' disease, primary glomerulonephritis
CC   and inflammatory bowel disease. The peptides, along with a detection
CC   agent, may additionally be used in a diagnostic composition. The peptides
CC   of the invention have a specific effect on autoreactive T-cells whilst
CC   leaving other components of the immune system intact. Therefore, the
CC   peptides should be safe, having no side effects. The present sequence
CC   represents residues 263-275 of human cartilage glycoprotein-39 (HC gp-39)
CC   which was used in an exemplification of the present invention.
XX
SQ   Sequence   13 AA;

  Query Match             100.0%;  Score 43;  DB 21;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 0.0083;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTLASSETG 9
              |||||||||
Db          3 FTLASSETG 11


RESULT 12
AAU03045
ID   AAU03045 standard; peptide; 13 AA.
XX
AC   AAU03045;
XX
DT   12-SEP-2001  (first entry)
XX
DE   Human cartilage gp-39 autoimmune peptide #1.
XX
KW   Human; autoimmune disease; epitope; antigen; rheumatoid arthritis;
KW   immunosuppression.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "Deaminated"
XX
PN   WO200129081-A1.
XX
PD   26-APR-2001.
XX
PF   12-OCT-2000; 2000WO-EP10230.
XX
PR   18-OCT-1999;   99EP-0203427.
XX
PA   (ALKU ) AKZO NOBEL NV.
XX
PI   Van Staveren CJ,  Timmers CM,  Van Galen PJM,  Knegtel RMA,  Boots AMH;
PI   Miltenburg AMM;
XX
DR   WPI; 2001-367293/38.
XX
PT   Modified peptides derived from human cartilage gp-39 protein epitope
PT   are useful in immunotherapy, especially in the treatment of arthritis
PT   -
XX
PS   Claim 1; Page 36; 52pp; English.
XX
CC   The sequence represents Human cartilage gp-39 protein epitope (amino
CC   acids 263-275) which is used as the template for making modified
CC   peptides for use in immunotherapy. The peptides are useful for inducing
CC   tolerance induction in patients suffering from autoimmune diseases such
CC   as rheumatoid arthritis. The peptides provide antigen-specific, non-toxic
CC   immunosuppression therapy as a very attractive alternative to
CC   non-specific immunosuppression.
XX
SQ   Sequence   13 AA;

  Query Match             100.0%;  Score 43;  DB 22;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 0.0083;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTLASSETG 9
```

```
              |||||||.|
Db          3 FTLASSETG 11


RESULT 13
AAU03046
ID   AAU03046 standard; peptide; 13 AA.
XX
AC   AAU03046;
XX
DT   12-SEP-2001  (first entry)
XX
DE   Human cartilage gp-39 autoimmune peptide #2.
XX
KW   Human; autoimmune disease; epitope; antigen; rheumatoid arthritis;
KW   immunosuppression.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /label= OTHER
FT                   /note= "Other= desaminoarginiinyl"
FT   Modified-site   13
FT                   /note= "C-terminus is amidated"
XX
PN   WO200129081-A1.
XX
PD   26-APR-2001.
XX
PF   12-OCT-2000; 2000WO-EP10230.
XX
PR   18-OCT-1999;   99EP-0203427.
XX
PA   (ALKU ) AKZO NOBEL NV.
XX
PI   Van Staveren CJ,  Timmers CM,  Van Galen PJM,  Knegtel RMA,  Boots AMH;
PI   Miltenburg AMM;
XX
DR   WPI; 2001-367293/38.
XX
PT   Modified peptides derived from human cartilage gp-39 protein epitope
PT   are useful in immunotherapy, especially in the treatment of arthritis
PT   -
XX
PS   Claim 12; Page 39; 52pp; English.
XX
CC   The sequence represents a modified peptide derived from Human cartilage
CC   gp-39 protein epitope (amino acids 263-275) for use in immunotherapy.
CC   The peptides are useful for inducing tolerance induction in patients
CC   suffering from autoimmune diseases such as rheumatoid arthritis.
CC   The peptides provide antigen-specific, non-toxic immunosuppression
CC   therapy as a very attractive alternative to non-specific
CC   immunosuppression.
XX
SQ   Sequence   13 AA;

  Query Match             100.0%;  Score 43;  DB 22;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 0.0083;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTLASSETG 9
              |||||||||
Db          3 FTLASSETG 11


RESULT 14
AAU03047
ID   AAU03047 standard; peptide; 13 AA.
XX
AC   AAU03047;
XX
DT   12-SEP-2001  (first entry)
XX
DE   Human cartilage gp-39 autoimmune peptide #3.
XX
KW   Human; autoimmune disease; epitope; antigen; rheumatoid arthritis;
KW   immunosuppression.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /label= OTHER
FT                   /note= "Other= desaminoarginiinyl"
XX
PN   WO200129081-A1.
XX
PD   26-APR-2001.
XX
PF   12-OCT-2000; 2000WO-EP10230.
XX
PR   18-OCT-1999;   99EP-0203427.
XX
PA   (ALKU ) AKZO NOBEL NV.
XX
PI   Van Staveren CJ,  Timmers CM,  Van Galen PJM,  Knegtel RMA,  Boots AMH;
PI   Miltenburg AMM;
XX
DR   WPI; 2001-367293/38.
XX
PT   Modified peptides derived from human cartilage gp-39 protein epitope
PT   are useful in immunotherapy, especially in the treatment of arthritis
PT   -
XX
PS   Claim 12; Page 39; 52pp; English.
XX
CC   The sequence represents a modified peptide derived from Human cartilage
CC   gp-39 protein epitope (amino acids 263-275) for use in immunotherapy.
CC   The peptides are useful for inducing tolerance induction in patients
CC   suffering from autoimmune diseases such as rheumatoid arthritis.
CC   The peptides provide antigen-specific, non-toxic immunosuppression
CC   therapy as a very attractive alternative to non-specific
CC   immunosuppression..
XX
SQ   Sequence   13 AA;

  Query Match             100.0%;  Score 43;  DB 22;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 0.0083;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTLASSETG 9
              |||||||||
Db          3 FTLASSETG 11


RESULT 15
AAU03048
ID   AAU03048 standard; peptide; 13 AA.
XX
AC   AAU03048;
XX
DT   12-SEP-2001  (first entry)
XX
DE   Human cartilage gp-39 autoimmune peptide #4.
XX
KW   Human; autoimmune disease; epitope; antigen; rheumatoid arthritis;
KW   immunosuppression.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
```

```
FT                   /note= "Connected to CH3-(OCH2CH3)3-OCH2C(O)"
FT   Modified-site   13
FT                   /note= "C-terminus is amidated"
XX
PN   WO200129081-A1.
XX
PD   26-APR-2001.
XX
PF   12-OCT-2000; 2000WO-EP10230.
XX
PR   18-OCT-1999;   99EP-0203427.
XX
PA   (ALKU ) AKZO NOBEL NV.
XX
PI   Van Staveren CJ,  Timmers CM,  Van Galen PJM,  Kneqtel RMA,  Boots AMH;
PI   Miltenburg AMM;
XX
DR   WPI; 2001-367293/38.
XX
PT   Modified peptides derived from human cartilage gp-39 protein epitope
PT   are useful in immunotherapy, especially in the treatment of arthritis
PT   -
XX
PS   Claim 12; Page 39; 52pp; English.
XX
CC   The sequence represents a modified peptide derived from Human cartilage
CC   gp-39 protein epitope (amino acids 263-275) for use in immunotherapy.
CC   The peptides are useful for inducing tolerance induction in patients
CC   suffering from autoimmune diseases such as rheumatoid arthritis.
CC   The peptides provide antigen-specific, non-toxic immunosuppression
CC   therapy as a very attractive alternative to non-specific
CC   immunosuppression.
XX
SQ   Sequence   13 AA;

  Query Match             100.0%;  Score 43;  DB 22;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 0.0093;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTLASSETG 9
              |||||||||
Db          3 FTLASSETG 11


Search completed: September 24, 2003, 13:41:20
Job time : 37.2261 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 24, 2003, 13:35:23 ; Search time 10.9213 Seconds
(without alignments)
34.867 Million cell updates/sec

Title: US-09-744-282-2
Perfect score: 43
Sequence: 1 FTLASSETG 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	43	100.0	9	1	US-08-619-645-2	Sequence 2, Appli
2	43	100.0	9	2	US-08-634-493-2	Sequence 2, Appli
3	43	100.0	13	1	US-08-619-645-6	Sequence 6, Appli
4	43	100.0	13	2	US-08-634-493-6	Sequence 6, Appli
5	43	100.0	13	4	US-09-405-745-1	Sequence 1, Appli
6	43	100.0	15	4	US-09-405-745-2	Sequence 2, Appli
7	43	100.0	16	3	US-09-171-705-41	Sequence 41, Appl
8	43	100.0	16	4	US-09-405-745-3	Sequence 3, Appli
9	35	81.4	481	4	US-08-976-063E-28	Sequence 28, Appl
10	34	79.1	385	2	US-08-694-915-2	Sequence 2, Appli
11	34	79.1	416	2	US-08-694-915-4	Sequence 4, Appli
12	34	79.1	423	4	US-08-850-348A-2	Sequence 2, Appli
13	33	76.7	13	4	US-09-459-749D-5	Sequence 5, Appli
14	33	76.7	125	1	US-08-478-039-69	Sequence 69, Appl
15	33	76.7	125	1	US-08-476-349A-69	Sequence 69, Appl
16	33	76.7	383	4	US-09-459-749D-17	Sequence 17, Appl
17	32	74.4	13	1	US-08-619-645-5	Sequence 5, Appli
18	32	74.4	13	2	US-08-634-493-5	Sequence 5, Appli
19	32	74.4	16	3	US-09-171-705-40	Sequence 40, Appl
20	32	74.4	502	3	US-09-413-814-106	Sequence 106, App
21	31	72.1	389	4	US-09-252-991A-31252	Sequence 31252, A
22	31	72.1	601	2	US-08-606-288-7	Sequence 7, Appli
23	31	72.1	601	2	US-08-606-288-10	Sequence 10, Appl
24	31	72.1	601	3	US-09-347-483-7	Sequence 7, Appli
25	31	72.1	601	3	US-09-347-483-10	Sequence 10, Appl
26	31	72.1	637	4	US-09-360-545-69	Sequence 69, Appl
27	30	69.8	263	2	US-08-860-577-4	Sequence 4, Appli
28	30	69.8	323	3	US-08-846-826A-2	Sequence 2, Appli
29	30	69.8	324	2	US-08-860-577-8	Sequence 8, Appli
30	30	69.8	607	4	US-08-556-422A-4	Sequence 4, Appli
31	29	67.4	178	3	US-09-000-630C-23	Sequence 23, Appl
32	29	67.4	178	3	US-08-862-730C-23	Sequence 23, Appl
33	29	67.4	178	3	US-09-417-455-10	Sequence 10, Appl
34	29	67.4	178	4	US-09-346-942-10	Sequence 10, Appl
35	29	67.4	178	4	US-09-316-081-6	Sequence 6, Appli
36	29	67.4	178	4	US-09-578-458-6	Sequence 6, Appli
37	29	67.4	178	4	US-09-522-964A-6	Sequence 6, Appli
38	29	67.4	178	4	US-09-457-626-10	Sequence 10, Appl
39	29	67.4	178	4	US-09-576-008-10	Sequence 10, Appl
40	29	67.4	401	2	US-08-591-079-4	Sequence 4, Appli
41	29	67.4	401	2	US-08-591-079-6	Sequence 6, Appli
42	29	67.4	457	4	US-09-252-991A-19877	Sequence 19877, A
43	29	67.4	458	4	US-09-428-352-7272	Sequence 7272, Ap
44	29	67.4	642	2	US-08-245-511-48	Sequence 48, Appl
45	29	67.4	642	2	US-08-600-993A-48	Sequence 48, Appl

ALIGNMENTS

```
RESULT 1
US-08-619-645-2
; Sequence 2, Application US/08619645
; Patent No. 5736507
; GENERAL INFORMATION:
;   APPLICANT: Boots, Anna M.H.
;   APPLICANT: Verheijden, Gijsbertus F.M.
;   TITLE OF INVENTION: No. 5736507el peptides derived from
;   TITLE OF INVENTION: autoantigen for use in immunotherapy of autoimmune
;   TITLE OF INVENTION: disease
;   NUMBER OF SEQUENCES: 9
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Akzo No. 5736507el Patent Department
;     STREET: 1300 Piccard Drive, Suite 206
;     CITY: Rockville
;     STATE: Maryland
;     COUNTRY: U.S.A.
;     ZIP: 20850
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.25(EPO)
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/619,645
;     FILING DATE: 25-MAR-1996
;     CLASSIFICATION: 514
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Gormley, Mary E.
;     REGISTRATION NUMBER: 34,409
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (301) 947-4433
; INFORMATION FOR SEQ ID NO: 2:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 9 amino acids
;     TYPE: amino acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;   MOLECULE TYPE: peptide
US-08-619-645-2

  Query Match          100.0%;  Score 43;  DB 1;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 2.5e-05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTLASSETG 9
              |||||||||
Db          1 FTLASSETG 9
```

```
RESULT 2
US-08-634-493-2
; Sequence 2, Application US/08634493
; Patent No. 5843449
;  GENERAL INFORMATION:
;    APPLICANT:  A.M.H. Boots
;    APPLICANT:  G.F.M. Verheijden
;    APPLICANT:  E.S. Bos
;    TITLE OF INVENTION:  No. 5843449el Peptides derived from autoantigen for use
;    TITLE OF INVENTION:  in Immunotherapy of Autoimmune Diseases
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Akzo No. 5843449el Patent Department
;      STREET:  1300 Piccard Drive, Suite 206
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/634,493
;      FILING DATE:  18-APR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/619,645
;      FILING DATE:  25-MAR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/EP95/04201
;      FILING DATE:  25-OCT-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  NL 942031287
;      FILING DATE:  27-OCT-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  NL 952008860
;      FILING DATE:  07-APR-1995
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 258-5200
;      TELEFAX:  (301) 977-0847
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-634-493-2

  Query Match             100.0%;  Score 43;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.5e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTLASSETG 9
              |||||||||
Db          1 FTLASSETG 9


RESULT 3
US-08-619-645-6
; Sequence 6, Application US/08619645
; Patent No. 5736507
;  GENERAL INFORMATION:
;    APPLICANT:  Boots, Anna M.H.
;    APPLICANT:  Verheijden, Gijsbertus F.M.
;    TITLE OF INVENTION:  No. 5736507el peptides derived from
;    TITLE OF INVENTION:  autoantigen for use in immunotherapy of autoimmune
;    TITLE OF INVENTION:  disease
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Akzo No. 5736507el Patent Department
;      STREET:  1300 Piccard Drive, Suite 206
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  U.S.A.
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25(EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/619,645
;      FILING DATE:  25-MAR-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Gormley, Mary E.
;      REGISTRATION NUMBER:  34,409
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 947-4433
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  13 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-619-645-6

  Query Match             100.0%;  Score 43;  DB 1;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 0.0029;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTLASSETG 9
              |||||||||
Db          3 FTLASSETG 11


RESULT 4
US-08-634-493-6
; Sequence 6, Application US/08634493
; Patent No. 5843449
;  GENERAL INFORMATION:
;    APPLICANT:  A.M.H. Boots
;    APPLICANT:  G.F.M. Verheijden
;    APPLICANT:  E.S. Bos
;    TITLE OF INVENTION:  No. 5843449el Peptides derived from autoantigen for use
;    TITLE OF INVENTION:  in Immunotherapy of Autoimmune Diseases
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Akzo No. 5843449el Patent Department
;      STREET:  1300 Piccard Drive, Suite 206
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/634,493
;      FILING DATE:  18-APR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/619,645
;      FILING DATE:  25-MAR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/EP95/04201
;      FILING DATE:  25-OCT-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  NL 942031287
;      FILING DATE:  27-OCT-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  NL 952008860
```

```
;      FILING DATE:  07-APR-1995
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 258-5200
;      TELEFAX:  (301) 977-0847
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  13 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-634-493-6

  Query Match           100.0%;  Score 43;  DB 2;  Length 13;
  Best Local Similarity  100.0%;  Pred. No. 0.0029;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTLASSETG 9
              |||||||||
Db          3 FTLASSETG 11


RESULT 5
US-09-405-745-1
; Sequence 1, Application US/09405745
; Patent No. 6392020
;  GENERAL INFORMATION:
;    APPLICANT:  Steenbakkers PGA
;    TITLE OF INVENTION:  Method of preparing a monoclonal
;    TITLE OF INVENTION:  antibody (as amended)
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Akzo No. 6392020el Patent Dept.
;      STREET:  1300 Piccard Drive, Suite 206
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  US
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/405,745
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/985,898
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sullivan, Michael G.
;      REGISTRATION NUMBER:  35,377
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  301-948-7400
;      TELEFAX:  301-948-9751
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  13 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FRAGMENT TYPE:  internal
US-09-405-745-1

  Query Match           100.0%;  Score 43;  DB 4;  Length 13;
  Best Local Similarity  100.0%;  Pred. No. 0.0029;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTLASSETG 9
              |||||||||
Db          3 FTLASSETG 11


RESULT 6
US-09-405-745-2
; Sequence 2, Application US/09405745
; Patent No. 6392020
;  GENERAL INFORMATION:
;    APPLICANT:  Steenbakkers PGA
;    TITLE OF INVENTION:  Method of preparing a monoclonal
;    TITLE OF INVENTION:  antibody (as amended)
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Akzo No. 6392020el Patent Dept.
;      STREET:  1300 Piccard Drive, Suite 206
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  US
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/405,745
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/985,898
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sullivan, Michael G.
;      REGISTRATION NUMBER:  35,377
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  301-948-7400
;      TELEFAX:  301-948-9751
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  15 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FRAGMENT TYPE:  internal
US-09-405-745-2

  Query Match           100.0%;  Score 43;  DB 4;  Length 15;
  Best Local Similarity  100.0%;  Pred. No. 0.0034;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTLASSETG 9
              |||||||||
Db          5 FTLASSETG 13


RESULT 7
US-09-171-705-41
; Sequence 41, Application US/09171705
; Patent No. 6184204
; GENERAL INFORMATION:
;  APPLICANT: BOOTS, ANNA M.H.
;  APPLICANT: VERHEIJDEN, GILBERTUS F.M.
;  TITLE OF INVENTION: NOVEL PEPTIDES SUITABLE FOR USE IN ANTIGEN SPECIFIC
;  TITLE OF INVENTION: IMMUNOSUPPRESSIVE THERAPY
;  FILE REFERENCE: O/96198 US
;  CURRENT APPLICATION NUMBER: US/09/171,705
;  CURRENT FILING DATE: 1999-02-09
;  NUMBER OF SEQ ID NOS: 78
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 41
;   LENGTH: 16
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
```

```
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: DERIVED FROM
;   OTHER INFORMATION: SEQUENCE OF HUMAN CARTILAGE (HC) gp39 PROTEIN
US-09-171-705-41

  Query Match             100.0%;  Score 43;  DB 3;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 0.0037;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTLASSETG 9
              |||||||||
Db          4 FTLASSETG 12


RESULT 8
US-09-405-745-3
; Sequence 3, Application US/09405745
; Patent No. 6392020
;  GENERAL INFORMATION:
;    APPLICANT:  Steenbakkers PGA
;    TITLE OF INVENTION:  Method of preparing a monoclonal
;    TITLE OF INVENTION:  antibody (as amended)
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Akzo No. 6392020el Patent Dept.
;      STREET:  1300 Piccard Drive, Suite 206
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  US
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/405,745
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/985,898
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sullivan, Michael G.
;      REGISTRATION NUMBER:  35,377
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  301-948-7400
;      TELEFAX:  301-948-9751
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  16 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FRAGMENT TYPE:  internal
US-09-405-745-3

  Query Match             100.0%;  Score 43;  DB 4;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 0.0037;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTLASSETG 9
              |||||||||
Db          4 FTLASSETG 12


RESULT 9
US-08-976-063E-28
; Sequence 28, Application US/08976063E
; Patent No. 6524831
; GENERAL INFORMATION:
;  APPLICANT:  Steinbuchel, Alexander
;  APPLICANT:  Priefert, Horst
;  APPLICANT:  Rabenhorst, Jurgen
;  TITLE OF INVENTION:  SYNTHETIC ENZYMES FOR THE PRODUCTION OF CONIFERYL
;  TITLE OF INVENTION:  ALCOHOL, CONIFERYLALDEHYDE, FERULIC ACID, VANILLIN AND
;  TITLE OF INVENTION:  VANILLIC ACID AND THEIR USE
;  FILE REFERENCE:  Bayer-9998-CAC
;  CURRENT APPLICATION NUMBER: US/08/976,063E
;  CURRENT FILING DATE:  1997-11-21
;  PRIOR APPLICATION NUMBER:  196 49 655.1 GERMANY
;  PRIOR FILING DATE:  1996-11-29
;  NUMBER OF SEQ ID NOS:  45
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 28
;    LENGTH: 481
;    TYPE: PRT
;    ORGANISM: not required under old rule
US-08-976-063E-28

  Query Match             81.4%;  Score 35;  DB 4;  Length 481;
  Best Local Similarity   77.8%;  Pred. No. 13;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FTLASSETG 9
              || |:||||
Db         83 FTAAASEIG 91


RESULT 10
US-08-694-915-2
; Sequence 2, Application US/08694915
; Patent No. 5811535
;  GENERAL INFORMATION:
;    APPLICANT:  Adamou, Julie
;    APPLICANT:  Kirkpatrick, Robert
;    APPLICANT:  Rosenberg, Martin
;    TITLE OF INVENTION:  HUMAN CARTILEGE GP39-LIKE GENE
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SmithKline Beecham Corporation
;      STREET:  709 Swedeland Road
;      CITY:  King of Prussia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19406-2799
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/694,915
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Han, William T
;      REGISTRATION NUMBER:  34,344
;      REFERENCE/DOCKET NUMBER:  ATG50017
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  610-270-5219
;      TELEFAX:  610-270-5090
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  385 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
```

```
;    FRAGMENT TYPE:  N-terminal
;    ORIGINAL SOURCE:
US-08-694-915-2

  Query Match             79.1%;  Score 34;  DB 2;  Length 385;
  Best Local Similarity   87.5%;  Pred. No. 16;
  Matches    7;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTLASSET 8
              ||:|:||
Db        268 FTLASAET 275


RESULT 11
US-08-694-915-4
; Sequence 4, Application US/08694915
; Patent No. 5811535
;  GENERAL INFORMATION:
;    APPLICANT:  Adamou, Julie
;    APPLICANT:  Kirkpatrick, Robert
;    APPLICANT:  Rosenberg, Martin
;    TITLE OF INVENTION:  HUMAN CARTILEGE GP39-LIKE GENE
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SmithKline Beecham Corporation
;      STREET:  709 Swedeland Road
;      CITY:  King of Prussia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19406-2799
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/694,915
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Han, William T
;      REGISTRATION NUMBER:  34,344
;      REFERENCE/DOCKET NUMBER:  ATG50017
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  610-270-5219
;      TELEFAX:  610-270-5090
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  416 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:  N-terminal
;    ORIGINAL SOURCE:
US-08-694-915-4

  Query Match             79.1%;  Score 34;  DB 2;  Length 416;
  Best Local Similarity   87.5%;  Pred. No. 16;
  Matches    7;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTLASSET 8
              |||||:||
Db        299 FTLASAET 306


RESULT 12
US-08-850-348A-2
; Sequence 2, Application US/08850348A
; Patent No. 6576427
; GENERAL INFORMATION:
;  APPLICANT: Kirkpatrick, et al.
;  TITLE OF INVENTION: Human Cartilage Glycoprotein
;  FILE REFERENCE: PF498
;  CURRENT APPLICATION NUMBER: US/08/850,348A
;  CURRENT FILING DATE:  1997-05-02
;  PRIOR APPLICATION NUMBER: 60/016,532
;  PRIOR FILING DATE: 1996-05-03
;  NUMBER OF SEQ ID NOS: 3
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
;   LENGTH: 423
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-08-850-348A-2

  Query Match             79.1%;  Score 34;  DB 4;  Length 423;
  Best Local Similarity   87.5%;  Pred. No. 18;
  Matches    7;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTLASSET 8
              |||||:||
Db        306 FTLASAET 313


RESULT 13
US-09-459-749D-5
; Sequence 5, Application US/09459749D
; Patent No. 6464975
; GENERAL INFORMATION:
;  APPLICANT: Millis, Albert J. T.
;  TITLE OF INVENTION: Compositions and Methods For Altering Cell Migration
;  FILE REFERENCE: 0794.016A
;  CURRENT APPLICATION NUMBER: US/09/459,749D
;  CURRENT FILING DATE:  1999-12-10
;  PRIOR APPLICATION NUMBER: 60/111,856
;  PRIOR FILING DATE: 1998-12-11
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
;   LENGTH: 13
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:SITE
;   NAME/KEY: SITE
;   LOCATION: (1)..(13)
;   OTHER INFORMATION: Peptide corresponding to amino acids 261- 273 of
;   OTHER INFORMATION: gp38k protein
US-09-459-749D-5

  Query Match             76.7%;  Score 33;  DB 4;  Length 13;
  Best Local Similarity   87.5%;  Pred. No. 3.47;
  Matches    7;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTLASSET 8
              |||||:||
Db          5 FTLASSKT 12


RESULT 14
US-08-478-039-69
; Sequence 69, Application US/08478039
; Patent No. 5681722
;  GENERAL INFORMATION:
;    APPLICANT:  Newman, Roland A.
;    APPLICANT:  Hanna, Nabil
;    APPLICANT:  Raab, Ronald W.
;    TITLE OF INVENTION:  Recombinant Antibodies for Human Therapy
;    NUMBER OF SEQUENCES:  114
```

```
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BURNS, DOANE, SWECKER & MATHIS
;      STREET:  699 Prince St.
;      CITY:  Alexandria
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22313-1404
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/478,039
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/379,072
;      FILING DATE:  25-JAN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/912,292
;      FILING DATE:  10-JUL-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/856,281
;      FILING DATE:  23-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/735,064
;      FILING DATE:  25-JUL-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Teskin Esq., Robin L.
;      REGISTRATION NUMBER:  35,030
;      REFERENCE/DOCKET NUMBER:  012712-160
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6620
;      TELEFAX:  703-836-2021
;  INFORMATION FOR SEQ ID NO:  69:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  125 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  not relevant
;    MOLECULE TYPE:  peptide
;    ORIGINAL SOURCE:
;      ORGANISM:  Monkey
;    POSITION IN GENOME:
;      CHROMOSOME/SEGMENT:  VH2 clone 2-10
US-08-478-039-69

  Query Match             76.7%;  Score 33;  DB 1;  Length 125;
  Best Local Similarity   66.7%;  Pred. No. 7;
  Matches    6;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTLASSETG 9
              |:|::||||
Db         27 FSLSTSETG 35


RESULT 15
US-08-476-349A-69
; Sequence 69, Application US/08476349A
; Patent No. 5750105
;  GENERAL INFORMATION:
;    APPLICANT:  Newman, Roland A.
;    APPLICANT:  Hanna, Nabil
;    APPLICANT:  Raab, Ronald W.
;    TITLE OF INVENTION:  Recombinant Antibodies for Human Therapy
;    NUMBER OF SEQUENCES:  114
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BURNS, DOANE, SWECKER & MATHIS
;      STREET:  699 Prince St.
;      CITY:  Alexandria
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22313-1404
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/476,349A
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/379,072
;      FILING DATE:  25-JAN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/912,292
;      FILING DATE:  10-JUL-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/856,281
;      FILING DATE:  23-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/735,064
;      FILING DATE:  25-JUL-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Teskin Esq., Robin L.
;      REGISTRATION NUMBER:  35,030
;      REFERENCE/DOCKET NUMBER:  012712-161
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6620
;      TELEFAX:  703-836-2021
;  INFORMATION FOR SEQ ID NO:  69:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  125 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  not relevant
;    MOLECULE TYPE:  peptide
;    ORIGINAL SOURCE:
;      ORGANISM:  Monkey
;    POSITION IN GENOME:
;      CHROMOSOME/SEGMENT:  VH2 clone 2-10
US-08-476-349A-69

  Query Match             76.7%;  Score 33;  DB 1;  Length 125;
  Best Local Similarity   66.7%;  Pred. No. 7;
  Matches    6;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTLASSETG 9
              |:|::||||
Db         27 FSLSTSEIG 35


Search completed: September 24, 2003, 13:51:13
Job time : 11.9213 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 24, 2003, 13:47:40 ; Search time 20.1236 Seconds
(without alignments)
67.670 Million cell updates/sec

Title: US-09-744-282-2
Perfect score: 43
Sequence: 1 PTLASSETG 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 566894 seqs, 151307093 residues

Total number of hits satisfying chosen parameters: 566894

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	43	100.0	13	10	US-09-985-065-1	Sequence 1, Appli
2	43	100.0	15	10	US-09-985-065-2	Sequence 2, Appli
3	43	100.0	16	10	US-09-985-065-3	Sequence 3, Appli
4	43	100.0	199	9	US-09-815-837-56	Sequence 56, Appl
5	43	100.0	199	9	US-09-815-837-60	Sequence 60, Appl
6	43	100.0	208	9	US-09-815-837-102	Sequence 102, App
7	43	100.0	209	9	US-09-815-837-58	Sequence 58, Appl
8	43	100.0	210	9	US-09-815-837-40	Sequence 40, Appl
9	43	100.0	211	9	US-09-815-837-39	Sequence 39, Appl
10	43	100.0	211	9	US-09-815-837-59	Sequence 59, Appl
11	43	100.0	212	9	US-09-815-837-37	Sequence 37, Appl
12	43	100.0	212	9	US-09-815-837-38	Sequence 38, Appl
13	43	100.0	213	9	US-09-815-837-41	Sequence 41, Appl
14	43	100.0	213	9	US-09-815-837-43	Sequence 43, Appl
15	43	100.0	213	9	US-09-815-837-44	Sequence 44, Appl
16	43	100.0	213	9	US-09-815-837-57	Sequence 57, Appl
17	43	100.0	214	9	US-09-815-837-42	Sequence 42, Appl
18	43	100.0	216	9	US-09-815-837-45	Sequence 45, Appl
19	43	100.0	216	9	US-09-815-837-46	Sequence 46, Appl
20	43	100.0	220	9	US-09-815-837-47	Sequence 47, Appl
21	43	100.0	220	9	US-09-815-837-48	Sequence 48, Appl
22	43	100.0	383	15	US-10-097-340-45	Sequence 45, Appl
23	35	81.4	481	8	US-08-976-063C-28	Sequence 28, Appl
24	33	76.7	13	10	US-09-459-749D-5	Sequence 5, Appli
25	33	76.7	383	10	US-09-459-749D-17	Sequence 17, Appl
26	31	72.1	453	15	US-10-128-714-8161	Sequence 8161, Ap
27	31	72.1	453	15	US-10-128-714-8161	Sequence 8161, Ap
28	31	72.1	459	12	US-10-224-999A-3470	Sequence 3470, Ap
29	31	72.1	637	12	US-10-625-145A-69	Sequence 69, Appl
30	31	72.1	637	14	US-10-041-007-28	Sequence 28, Appl
31	31	72.1	4809	15	US-10-156-761-9090	Sequence 9090, Ap
32	30	69.8	75	12	US-10-238-075-1195	Sequence 1195, Ap
33	30	69.8	433	10	US-09-738-626-6405	Sequence 6405, Ap
34	30	69.8	439	15	US-10-156-761-11643	Sequence 11643, A
35	30	69.8	607	12	US-10-320-769-4	Sequence 4, Appli
36	30	69.8	782	10	US-09-908-193-47	Sequence 47, Appl
37	30	69.8	1294	11	US-09-836-499-2	Sequence 2, Appli
38	30	69.8	1294	15	US-10-162-435-2	Sequence 2, Appli
39	30	69.8	1309	11	US-09-836-499-5	Sequence 5, Appli
40	30	69.8	1309	15	US-10-162-435-5	Sequence 5, Appli
41	29	67.4	87	15	US-10-072-602B-184	Sequence 184, App
42	29	67.4	88	15	US-10-072-602B-249	Sequence 249, App
43	29	67.4	91	15	US-10-072-602B-134	Sequence 134, App
44	29	67.4	92	15	US-10-072-602B-252	Sequence 252, App
45	29	67.4	93	15	US-10-072-602B-201	Sequence 201, App

ALIGNMENTS

```
RESULT 1
US-09-985-065-1
; Sequence 1, Application US/09985065
; Patent No. US20020143150A1
;  GENERAL INFORMATION:
;       APPLICANT: Steenbakkers PGA
;       TITLE OF INVENTION: Method of preparing a monoclonal
;                           antibody (as amended)
;       NUMBER OF SEQUENCES: 3
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Akzo No. US20020143150A1e1 Patent Dept.
;            STREET: 1300 Piccard Drive, Suite 206
;            CITY: Rockville
;            STATE: Maryland
;            COUNTRY: US
;            ZIP: 20850
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/985,065
;            FILING DATE: 01-No. US20020143150A1-2001
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 09/405,745
;            FILING DATE: 01-NOV-2001
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Sullivan, Michael G.
;            REGISTRATION NUMBER: 35,377
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 301-948-7400
;            TELEFAX: 301-948-9751
;  INFORMATION FOR SEQ ID NO: 1:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 13 amino acids
;            TYPE: amino acid
```

```
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;          MOLECULE TYPE: protein
;          FRAGMENT TYPE: internal
;          SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-985-065-1

  Query Match             100.0%;  Score 43;  DB 10;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 0.018;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTLASSETG 9
              |||||||||
Db          3 FTLASSETG 11


RESULT 2
US-09-985-065-2
; Sequence 2, Application US/09985065
; Patent No. US20020143150A1
;   GENERAL INFORMATION:
;          APPLICANT: Steenbakkers PGA
;          TITLE OF INVENTION: Method of preparing a monoclonal
;                              antibody (as amended)
;          NUMBER OF SEQUENCES: 3
;          CORRESPONDENCE ADDRESS:
;               ADDRESSEE: Akzo No. US20020143150A1e1 Patent Dept.
;               STREET: 1300 Piccard Drive, Suite 206
;               CITY: Rockville
;               STATE: Maryland
;               COUNTRY: US
;               ZIP: 20850
;          COMPUTER READABLE FORM:
;               MEDIUM TYPE: Floppy disk
;               COMPUTER: IBM PC compatible
;               OPERATING SYSTEM: PC-DOS/MS-DOS
;               SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
;          CURRENT APPLICATION DATA:
;               APPLICATION NUMBER: US/09/985,065
;               FILING DATE: 01-No. US20020143150A1-2001
;               CLASSIFICATION: <Unknown>
;          PRIOR APPLICATION DATA:
;               APPLICATION NUMBER: 09/405,745
;               FILING DATE: 01-NOV-2001
;          ATTORNEY/AGENT INFORMATION:
;               NAME: Sullivan, Michael G.
;               REGISTRATION NUMBER: 35,377
;          TELECOMMUNICATION INFORMATION:
;               TELEPHONE: 301-948-7400
;               TELEFAX: 301-948-9751
;     INFORMATION FOR SEQ ID NO: 2:
;          SEQUENCE CHARACTERISTICS:
;               LENGTH: 15 amino acids
;               TYPE: amino acid
;               STRANDEDNESS: single
;               TOPOLOGY: linear
;          MOLECULE TYPE: protein
;          FRAGMENT TYPE: internal
;          SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-985-065-2

  Query Match             100.0%;  Score 43;  DB 10;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 0.021;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTLASSETG 9
              |||||||||
Db          5 FTLASSETG 13


RESULT 3
US-09-985-065-3
; Sequence 3, Application US/09985065
; Patent No. US20020143150A1
;   GENERAL INFORMATION:
;          APPLICANT: Steenbakkers PGA
;          TITLE OF INVENTION: Method of preparing a monoclonal
;                              antibody (as amended)
;          NUMBER OF SEQUENCES: 3
;          CORRESPONDENCE ADDRESS:
;               ADDRESSEE: Akzo No. US20020143150A1e1 Patent Dept.
;               STREET: 1300 Piccard Drive, Suite 206
;               CITY: Rockville
;               STATE: Maryland
;               COUNTRY: US
;               ZIP: 20850
;          COMPUTER READABLE FORM:
;               MEDIUM TYPE: Floppy disk
;               COMPUTER: IBM PC compatible
;               OPERATING SYSTEM: PC-DOS/MS-DOS
;               SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
;          CURRENT APPLICATION DATA:
;               APPLICATION NUMBER: US/09/985,065
;               FILING DATE: 01-No. US20020143150A1-2001
;               CLASSIFICATION: <Unknown>
;          PRIOR APPLICATION DATA:
;               APPLICATION NUMBER: 09/405,745
;               FILING DATE: 01-NOV-2001
;          ATTORNEY/AGENT INFORMATION:
;               NAME: Sullivan, Michael G.
;               REGISTRATION NUMBER: 35,377
;          TELECOMMUNICATION INFORMATION:
;               TELEPHONE: 301-948-7400
;               TELEFAX: 301-948-9751
;     INFORMATION FOR SEQ ID NO: 3:
;          SEQUENCE CHARACTERISTICS:
;               LENGTH: 16 amino acids
;               TYPE: amino acid
;               STRANDEDNESS: single
;               TOPOLOGY: linear
;          MOLECULE TYPE: protein
;          FRAGMENT TYPE: internal
;          SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-985-065-3

  Query Match             100.0%;  Score 43;  DB 10;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 0.023;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTLASSETG 9
              |||||||||
Db          4 FTLASSETG 12


RESULT 4
US-09-815-837-56
; Sequence 56, Application US/09815837
; Patent No. US20020082411A1
; GENERAL INFORMATION:
;  APPLICANT: Carter, Darrick
;  APPLICANT:  Zhu, Shirley
;  APPLICANT:  Arimilli, Subhashini
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Immune Medators and Related Methods
;  FILE REFERENCE: 014058-005670US
;  CURRENT APPLICATION NUMBER: US/09/815,837
;  CURRENT FILING DATE:  2001-03-22
;  PRIOR APPLICATION NUMBER: US 60/191,274
;  PRIOR FILING DATE: 2000-03-22
;  PRIOR APPLICATION NUMBER: US 60/204,249
;  PRIOR FILING DATE: 2000-05-15
;  PRIOR APPLICATION NUMBER: US 60/264,003
;  PRIOR FILING DATE: 2001-01-23
;  NUMBER OF SEQ ID NOS: 129
;  SOFTWARE: PatentIn Ver. 2.1
```

```
; SEQ ID NO 56
;   LENGTH: 199
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:CO528-AC single
;   OTHER INFORMATION: chain recombinant MHC class II molecule
US-09-815-837-56

  Query Match           100.0%;  Score 43;  DB 9;  Length 199;
  Best Local Similarity  100.0%;  Pred. No. 0.38;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTLASSETG 9
              |||||||||
Db          8 FTLASSETG 16


RESULT 5
US-09-815-837-60
; Sequence 60, Application US/09815837
; Patent No. US20020082411A1
; GENERAL INFORMATION:
;  APPLICANT: Carter, Darrick
;  APPLICANT:  Zhu, Shirley
;  APPLICANT:  Arimilli, Subhashini
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Immune Medators and Related Methods
;  FILE REFERENCE: 014058-005670US
;  CURRENT APPLICATION NUMBER: US/09/815,837
;  CURRENT FILING DATE:  2001-03-22
;  PRIOR APPLICATION NUMBER: US 60/191,274
;  PRIOR FILING DATE: 2000-03-22
;  PRIOR APPLICATION NUMBER: US 60/204,249
;  PRIOR FILING DATE: 2000-05-15
;  PRIOR APPLICATION NUMBER: US 60/264,003
;  PRIOR FILING DATE: 2001-01-23
;  NUMBER OF SEQ ID NOS: 129
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 60
;   LENGTH: 199
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:CO528 single
;   OTHER INFORMATION: chain recombinant MHC class II molecule
US-09-815-837-60

  Query Match           100.0%;  Score 43;  DB 9;  Length 199;
  Best Local Similarity  100.0%;  Pred. No. 0.38;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTLASSETG 9
              |||||||||
Db          8 FTLASSETG 16


RESULT 6
US-09-815-837-102
; Sequence 102, Application US/09815837
; Patent No. US20020082411A1
; GENERAL INFORMATION:
;  APPLICANT: Carter, Darrick
;  APPLICANT:  Zhu, Shirley
;  APPLICANT:  Arimilli, Subhashini
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Immune Medators and Related Methods
;  FILE REFERENCE: 014058-005670US
;  CURRENT APPLICATION NUMBER: US/09/815,837
;  CURRENT FILING DATE:  2001-03-22
;  PRIOR APPLICATION NUMBER: US 60/191,274
;  PRIOR FILING DATE: 2000-03-22
;  PRIOR APPLICATION NUMBER: US 60/204,249
;  PRIOR FILING DATE: 2000-05-15
;  PRIOR APPLICATION NUMBER: US 60/264,003
;  PRIOR FILING DATE: 2001-01-23
;  NUMBER OF SEQ ID NOS: 129
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 102
;   LENGTH: 208
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial
;   OTHER INFORMATION: Sequence:gp39-beta1/alpha1 HLA-DR4 CO563 human
;   OTHER INFORMATION: single chain molecule
US-09-815-837-102

  Query Match           100.0%;  Score 43;  DB 9;  Length 208;
  Best Local Similarity  100.0%;  Pred. No. 0.4;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTLASSETG 9
              |||||||||
Db          8 FTLASSETG 16


RESULT 7
US-09-815-837-58
; Sequence 58, Application US/09815837
; Patent No. US20020082411A1
; GENERAL INFORMATION:
;  APPLICANT: Carter, Darrick
;  APPLICANT:  Zhu, Shirley
;  APPLICANT:  Arimilli, Subhashini
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Immune Medators and Related Methods
;  FILE REFERENCE: 014058-005670US
;  CURRENT APPLICATION NUMBER: US/09/815,837
;  CURRENT FILING DATE:  2001-03-22
;  PRIOR APPLICATION NUMBER: US 60/191,274
;  PRIOR FILING DATE: 2000-03-22
;  PRIOR APPLICATION NUMBER: US 60/204,249
;  PRIOR FILING DATE: 2000-05-15
;  PRIOR APPLICATION NUMBER: US 60/264,003
;  PRIOR FILING DATE: 2001-01-23
;  NUMBER OF SEQ ID NOS: 129
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 58
;   LENGTH: 209
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:CO608 single
;   OTHER INFORMATION: chain recombinant MHC class II molecule
US-09-815-837-58

  Query Match           100.0%;  Score 43;  DB 9;  Length 209;
  Best Local Similarity  100.0%;  Pred. No. 0.4;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTLASSETG 9
              |||||||||
Db          7 FTLASSETG 15


RESULT 8
US-09-815-837-40
; Sequence 40, Application US/09815837
; Patent No. US20020082411A1
; GENERAL INFORMATION:
;  APPLICANT: Carter, Darrick
;  APPLICANT:  Zhu, Shirley
```

```
;  APPLICANT:  Arimilli, Subhashini
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Immune Medators and Related Methods
;  FILE REFERENCE: 014058-005670US
;  CURRENT APPLICATION NUMBER: US/09/815,837
;  CURRENT FILING DATE:  2001-03-22
;  PRIOR APPLICATION NUMBER: US 60/191,274
;  PRIOR FILING DATE: 2000-03-22
;  PRIOR APPLICATION NUMBER: US 60/204,249
;  PRIOR FILING DATE: 2000-05-15
;  PRIOR APPLICATION NUMBER: US 60/264,003
;  PRIOR FILING DATE: 2001-01-23
;  NUMBER OF SEQ ID NOS: 129
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 40
;   LENGTH: 210
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:CO595 single
;   OTHER INFORMATION: chain recombinant MHC class II molecule
US-09-815-837-40

  Query Match           100.0%;  Score 43;  DB 9;  Length 210;
  Best Local Similarity  100.0%;  Pred. No. 0.4;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTLASSETG 9
              |||| ||||
Db          8 FTLASSETG 16


RESULT 9
US-09-815-837-39
; Sequence 39, Application US/09815837
; Patent No. US20020082411A1
; GENERAL INFORMATION:
;  APPLICANT: Carter, Darrick
;  APPLICANT:  Zhu, Shirley
;  APPLICANT:  Arimilli, Subhashini
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Immune Medators and Related Methods
;  FILE REFERENCE: 014058-005670US
;  CURRENT APPLICATION NUMBER: US/09/815,837
;  CURRENT FILING DATE:  2001-03-22
;  PRIOR APPLICATION NUMBER: US 60/191,274
;  PRIOR FILING DATE: 2000-03-22
;  PRIOR APPLICATION NUMBER: US 60/204,249
;  PRIOR FILING DATE: 2000-05-15
;  PRIOR APPLICATION NUMBER: US 60/264,003
;  PRIOR FILING DATE: 2001-01-23
;  NUMBER OF SEQ ID NOS: 129
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 39
;   LENGTH: 211
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:CO596 single
;   OTHER INFORMATION: chain recombinant MHC class II molecule
US-09-815-837-39

  Query Match           100.0%;  Score 43;  DB 9;  Length 211;
  Best Local Similarity  100.0%;  Pred. No. 0.4;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTLASSETG 9
              |||| ||||
Db          8 FTLASSETG 16


RESULT 10
US-09-815-837-59
; Sequence 59, Application US/09815837
; Patent No. US20020082411A1
; GENERAL INFORMATION:
;  APPLICANT: Carter, Darrick
;  APPLICANT:  Zhu, Shirley
;  APPLICANT:  Arimilli, Subhashini
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Immune Medators and Related Methods
;  FILE REFERENCE: 014058-005670US
;  CURRENT APPLICATION NUMBER: US/09/815,837
;  CURRENT FILING DATE:  2001-03-22
;  PRIOR APPLICATION NUMBER: US 60/191,274
;  PRIOR FILING DATE: 2000-03-22
;  PRIOR APPLICATION NUMBER: US 60/204,249
;  PRIOR FILING DATE: 2000-05-15
;  PRIOR APPLICATION NUMBER: US 60/264,003
;  PRIOR FILING DATE: 2001-01-23
;  NUMBER OF SEQ ID NOS: 129
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 59
;   LENGTH: 211
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:CO608 variation
;   OTHER INFORMATION: single chain recombinant MHC class II molecule
US-09-815-837-59

  Query Match           100.0%;  Score 43;  DB 9;  Length 211;
  Best Local Similarity  100.0%;  Pred. No. 0.4;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTLASSETG 9
              |||| ||||
Db          8 FTLASSETG 16


RESULT 11
US-09-815-837-37
; Sequence 37, Application US/09815837
; Patent No. US20020082411A1
; GENERAL INFORMATION:
;  APPLICANT: Carter, Darrick
;  APPLICANT:  Zhu, Shirley
;  APPLICANT:  Arimilli, Subhashini
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Immune Medators and Related Methods
;  FILE REFERENCE: 014058-005670US
;  CURRENT APPLICATION NUMBER: US/09/815,837
;  CURRENT FILING DATE:  2001-03-22
;  PRIOR APPLICATION NUMBER: US 60/191,274
;  PRIOR FILING DATE: 2000-03-22
;  PRIOR APPLICATION NUMBER: US 60/204,249
;  PRIOR FILING DATE: 2000-05-15
;  PRIOR APPLICATION NUMBER: US 60/264,003
;  PRIOR FILING DATE: 2001-01-23
;  NUMBER OF SEQ ID NOS: 129
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 37
;   LENGTH: 212
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:CO598 single
;   OTHER INFORMATION: chain recombinant MHC class II molecule, human MHC
;   OTHER INFORMATION: class II single chain molecule with murine linkers
US-09-815-837-37

  Query Match           100.0%;  Score 43;  DB 9;  Length 212;
```

```
  Best Local Similarity   100.0%;  Pred. No. 0.4;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTLASSETG 9
              |||||||||
Db          8 FTLASSETG 16


RESULT 12
US-09-815-837-38
; Sequence 38, Application US/09815837
; Patent No. US20020082411A1
; GENERAL INFORMATION:
;  APPLICANT: Carter, Darrick
;  APPLICANT:  Zhu, Shirley
;  APPLICANT:  Arimilli, Subhashini
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Immune Medators and Related Methods
;  FILE REFERENCE: 014058-005670US
;  CURRENT APPLICATION NUMBER: US/09/815,837
;  CURRENT FILING DATE:  2001-03-22
;  PRIOR APPLICATION NUMBER: US 60/191,274
;  PRIOR FILING DATE: 2000-03-22
;  PRIOR APPLICATION NUMBER: US 60/204,249
;  PRIOR FILING DATE: 2000-05-15
;  PRIOR APPLICATION NUMBER: US 60/264,003
;  PRIOR FILING DATE: 2001-01-23
;  NUMBER OF SEQ ID NOS: 129
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 38
;   LENGTH: 212
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:CO597 single
;   OTHER INFORMATION: chain recombinant MHC class II molecule
US-09-815-837-38

  Query Match             100.0%;  Score 43;  DB 9;  Length 212;
  Best Local Similarity   100.0%;  Pred. No. 0.4;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTLASSETG 9
              |||||||||
Db          8 FTLASSETG 16


RESULT 13
US-09-815-837-41
; Sequence 41, Application US/09815837
; Patent No. US20020082411A1
; GENERAL INFORMATION:
;  APPLICANT: Carter, Darrick
;  APPLICANT:  Zhu, Shirley
;  APPLICANT:  Arimilli, Subhashini
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Immune Medators and Related Methods
;  FILE REFERENCE: 014058-005670US
;  CURRENT APPLICATION NUMBER: US/09/815,837
;  CURRENT FILING DATE:  2001-03-22
;  PRIOR APPLICATION NUMBER: US 60/191,274
;  PRIOR FILING DATE: 2000-03-22
;  PRIOR APPLICATION NUMBER: US 60/204,249
;  PRIOR FILING DATE: 2000-05-15
;  PRIOR APPLICATION NUMBER: US 60/264,003
;  PRIOR FILING DATE: 2001-01-23
;  NUMBER OF SEQ ID NOS: 129
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41
;   LENGTH: 213
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:CO593 single
;   OTHER INFORMATION: chain recombinant MHC class II molecule
US-09-815-837-41

  Query Match             100.0%;  Score 43;  DB 9;  Length 213;
  Best Local Similarity   100.0%;  Pred. No. 0.41;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTLASSETG 9
              |||||||||
Db          8 FTLASSETG 16


RESULT 14
US-09-815-837-43
; Sequence 43, Application US/09815837
; Patent No. US20020082411A1
; GENERAL INFORMATION:
;  APPLICANT: Carter, Darrick
;  APPLICANT:  Zhu, Shirley
;  APPLICANT:  Arimilli, Subhashini
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Immune Medators and Related Methods
;  FILE REFERENCE: 014058-005670US
;  CURRENT APPLICATION NUMBER: US/09/815,837
;  CURRENT FILING DATE:  2001-03-22
;  PRIOR APPLICATION NUMBER: US 60/191,274
;  PRIOR FILING DATE: 2000-03-22
;  PRIOR APPLICATION NUMBER: US 60/204,249
;  PRIOR FILING DATE: 2000-05-15
;  PRIOR APPLICATION NUMBER: US 60/264,003
;  PRIOR FILING DATE: 2001-01-23
;  NUMBER OF SEQ ID NOS: 129
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 43
;   LENGTH: 213
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:CO584 single
;   OTHER INFORMATION: chain recombinant MHC class II molecule
US-09-815-837-43

  Query Match             100.0%;  Score 43;  DB 9;  Length 213;
  Best Local Similarity   100.0%;  Pred. No. 0.41;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTLASSETG 9
              |||||||||
Db          8 FTLASSETG 16


RESULT 15
US-09-815-837-44
; Sequence 44, Application US/09815837
; Patent No. US20020082411A1
; GENERAL INFORMATION:
;  APPLICANT: Carter, Darrick
;  APPLICANT:  Zhu, Shirley
;  APPLICANT:  Arimilli, Subhashini
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Immune Medators and Related Methods
;  FILE REFERENCE: 014058-005670US
;  CURRENT APPLICATION NUMBER: US/09/815,837
;  CURRENT FILING DATE:  2001-03-22
;  PRIOR APPLICATION NUMBER: US 60/191,274
;  PRIOR FILING DATE: 2000-03-22
;  PRIOR APPLICATION NUMBER: US 60/204,249
;  PRIOR FILING DATE: 2000-05-15
```

```
;  PRIOR APPLICATION NUMBER: US 60/264,003
;  PRIOR FILING DATE: 2001-01-23
;  NUMBER OF SEQ ID NOS: 129
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 44
;    LENGTH: 213
;    TYPE: PRT
;    ORGANISM: Artificial Sequence
;    FEATURE:
;    OTHER INFORMATION: Description of Artificial Sequence:CO583 single
;    OTHER INFORMATION: chain recombinant MHC class II molecule
US-09-815-837-44

  Query Match             100.0%;  Score 43;  DB 9;  Length 213;
  Best Local Similarity   100.0%;  Pred. No. 0.41;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTLASSETG 9
              |||||||||
Db          8 FTLASSETG 16


Search completed: September 24, 2003, 14:22:26
Job time : 21.1236 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 24, 2003, 13:34:08 ; Search time 10.921s Seconds
(without alignments)
79.250 Million cell updates/sec

Title: US-09-744-282-2
Perfect score: 43
Sequence: 1 FTLASSETG 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	43	100.0	383	2	A49562	cartilage glycopro
2	37	86.0	441	2	T31090	probable ATP-bindi
3	33	76.7	383	2	S51327	heparin-binding gl
4	32	74.4	251	2	C69046	conserved hypothet
5	32	74.4	261	2	F84325	hypothetical prote
6	32	74.4	360	2	S66117	yacK protein - Bac
7	32	74.4	405	2	S61551	breast-regressing
8	32	74.4	574	2	A84782	hypothetical prote
9	32	74.4	986	2	AB2209	two-component sens
10	31	72.1	187	2	S69542	hypothetical prote
11	31	72.1	202	2	G87353	hypothetical prote
12	31	72.1	318	2	T20178	hypothetical prote
13	31	72.1	320	2	T24965	hypothetical prote
14	31	72.1	324	2	T20681	hypothetical prote
15	31	72.1	373	2	D83331	muconate cycloisom
16	31	72.1	414	2	T43387	hypothetical prote
17	31	72.1	445	2	B40590	flagellar distal c
18	31	72.1	459	1	P2WU6	L2 protein - human
19	31	72.1	491	2	T20061	hypothetical prote
20	31	72.1	515	2	S38270	amidase (EC 3.5.1.
21	31	72.1	590	1	GB8P24	baseplate protein
22	31	72.1	605	2	E69153	cadmium efflux ATP
23	31	72.1	860	2	T49910	hypothetical prote
24	31	72.1	861	2	T49909	hypothetical prote
25	31	72.1	1277	2	T14152	synaptic scaffoldi
26	30	69.8	167	2	AB2796	acetyltransferase
27	30	69.8	167	2	B97575	hypothetical prote
28	30	69.8	241	2	S71473	endo-1,4-beta-xyla
29	30	69.8	246	2	A25779	retrovirus-related
30	30	69.8	266	2	C70758	hypothetical prote
31	30	69.8	317	2	S45708	MSH receptor - bov
32	30	69.8	346	2	E72672	hypothetical prote
33	30	69.8	365	2	AG1246	aminotripeptidase
34	30	69.8	375	2	AH0306	probable exported
35	30	69.8	433	2	S15471	pol polyprotein -
36	30	69.8	447	2	T06004	1-aminocyclopropan
37	30	69.8	461	2	T03414	probable 1-aminocy
38	30	69.8	488	2	D70876	probable polyketid
39	30	69.8	596	2	AC1318	two-component sens
40	30	69.8	782	2	I48746	semaphorin C - mou
41	30	69.8	893	2	T03864	hypothetical prote
42	29	67.4	163	2	S11149	amiB protein - Str
43	29	67.4	167	2	AG2292	hypothetical prote
44	29	67.4	178	2	C40956	interleukin-1 rece
45	29	67.4	197	2	I43066	hypothetical prote

ALIGNMENTS

RESULT 1
A49562
cartilage glycoprotein gp39 precursor - human
N;Alternate names: 39K synovial protein
C;Species: Homo sapiens (man)
C;Date: 23-Mar-1995 #sequence_revision 23-Mar-1995 #text_change 22-Jun-1999
C;Accession: A49562; S10677; A33162
R;Hakala, B.E.; White, C.; Recklies, A.D.
J. Biol. Chem. 268, 25803-25810, 1993
A;Title: Human cartilage gp-39, a major secretory product of articular chondrocytes
A;Reference number: A49562; MUID:94064658; PMID:8245017
A;Accession: A49562
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-383 <HAK>
A;Cross-references: GB:M80927; NID:g348911; PIDN:AAA16074.1; PID:g348912
R;Nyirkos, P.; Golds, E.E.
Biochem. J. 269, 265-268, 1990
A;Title: Human synovial cells secrete a 39 kDa protein similar to a bovine mammary
A;Reference number: S10677; MUID:90328983; PMID:2375755
A;Accession: S10677
A;Molecule type: protein
A;Residues: 22-40,'X',42-45 <NY2>
C;Superfamily: Streptomyces chitinase ch140
C;Keywords: cartilage; extracellular protein; glycoprotein
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-383/Product: cartilage glycoprotein gp39 #status predicted <MAT>

Query Match 100.0%; Score 43; DB 2; Length 383;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 FTLASSETG 9
              |||||||||
Db        265 FTLASSETG 273
```

RESULT 2
T31090
probable ATP-binding protein - Serratia marcescens
C;Species: Serratia marcescens
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 17-Mar-2000
C;Accession: T31090
R;Salgi, F.; Climent, N.; Pique, N.; Sanchez, C.; Merino, S.; Rubires, X.; Aguilar,
J. Bacteriol. 181, 1883-1891, 1999
A;Title: Genetic analysis of the Serratia marcescens N28b O4 antigen gene cluster.
A;Reference number: Z20974; MUID:99173913; PMID:10074083
A;Accession: T31090
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-441 <SAI>
A;Cross-references: EMBL:AF038816; NID:g2828669; PID:g2828672; PIDN:AAC00182.1

```
C;Genetics:
A;Gene: wzt
C;Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology

  Query Match           86.0%;  Score 37;  DB 2;  Length 441;
  Best Local Similarity  100.0%;  Pred. No. 2.9;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTLASSET 8
              ||||||||
Db        332 FTLASSET 339


RESULT 3
S51327
heparin-binding glycoprotein 38K - pig
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 22-Jun-1999
C;Accession: S51327
R;Shackelton, L.M.; Mann, D.M.; Millis, A.J.T.
submitted to the EMBL Data Library, January 1995
A;Description: Identification of a 38kDa heparin-binding glycoprotein (gp38k) in differe
A;Reference number: S51327
A;Accession: S51327
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-383 <SHA>
A;Cross-references: EMBL:Z47803; NID:g634097; PIDN:CAA87764.1; PID:g634096
C;Superfamily: Streptomyces chitinase chi40

  Query Match           76.7%;  Score 33;  DB 2;  Length 383;
  Best Local Similarity  87.5%;  Pred. No. 20;
  Matches    7;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTLASSET 8
              |||:|||:|
Db        265 FTLASSKT 272


RESULT 4
C69046
conserved hypothetical protein MTH1349 - Methanobacterium thermoautotrophicum (strain De
C;Species: Methanobacterium thermoautotrophicum
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
C;Accession: C69046
R;Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.
ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A;Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: func
A;Reference number: A69000; MUID:98037514; PMID:9371463
A;Accession: C69046
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-251 <MTH>
A;Cross-references: GB:AE000898; GB:AE000666; NID:g2622453; PIDN:AAB85826.1; PID:g262245
A;Experimental source: strain Delta H
C;Genetics:
A;Gene: MTH1349

  Query Match           74.4%;  Score 32;  DB 2;  Length 251;
  Best Local Similarity  66.7%;  Pred. No. 21;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FTLASSETG 9
              |:| |:|:|
Db          7 FTIFSTETG 15


RESULT 5
F84325
hypothetical protein Vng1735c [imported] - Halobacterium sp. NRC-1
C;Species: Halobacterium sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C;Accession: F84325
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lask
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; J
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.
A;Title: Genome sequence of Halobacterium species NRC-1.
A;Reference number: A84160; MUID:20504483; PMID:11016950
A;Accession: F84325
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-261 <STO>
A;Cross-references: GB:AE004437; NID:g10581197; PIDN:AAG19970.1; GSPDB:GN00138
C;Genetics:
A;Gene: VNG1735C

  Query Match           74.4%;  Score 32;  DB 2;  Length 261;
  Best Local Similarity  66.7%;  Pred. No. 22;
  Matches    6;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 FTLASSETG 9
              |||  |:||
Db         37 FTLVDSDTG 45


RESULT 6
S66117
yacK protein - Bacillus subtilis
C;Species: Bacillus subtilis
C;Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 20-Jun-2000
C;Accession: S66117; C69741
R;Ogasawara, N.; Nakai, S.; Yoshikawa, H.
DNA Res. 1, 1-14, 1994
A;Title: Systematic sequencing of the 180 kilobase region of the Bacillus subtilis c
A;Reference number: S65967; MUID:96051385; PMID:7584024
A;Accession: S66117
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-360 <OGA>
A;Cross-references: EMBL:D26185; NID:g467326; PIDN:BAA05322.1; PID:g467476
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, December 199
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Be
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari,
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Ga
Lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo,
Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardin
A;Authors: Lauber, J.; Nazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Ma
y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portet
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scan
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; S
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiy
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshid
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtil
A;Reference number: A69580; MUID:98044033; PMID:9384377
A;Accession: C69741
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-360 <KUN>
A;Cross-references: GB:Z99104; GB:AL009126; NID:g2632267; PIDN:CAB11864.1; PID:g2632
A;Experimental source: strain 168
C;Genetics:
A;Gene: yacK

  Query Match           74.4%;  Score 32;  DB 2;  Length 360;
  Best Local Similarity  75.0%;  Pred. No. 32;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 TLASSETG 9
              |::||||
```

```
Db        101 TISSSETG 108


RESULT 7
S61551
breast-regressing protein brp39 precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
C;Accession: S61551; S61550; I48271
R;Morrison, B.W.; Leder, P.
Oncogene 9, 3417-3426, 1994
A;Title: neu and ras initiate murine mammary tumors that share genetic markers generally
A;Reference number: I48271; MUID:95060797; PMID:7970705
A;Accession: S61551
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-405 <MOR1>
A;Cross-references: EMBL:X93035; NID:g1085065; PIDN:CAA63603.1; PID:g1085066
R;Morrison, B.W.
submitted to the EMBL Data Library, November 1995
A;Reference number: S61550
A;Accession: S61550
A;Molecule type: mRNA
A;Residues: 1-245,'S',247-330,'H',332-350,'MVWALDLDDPQGTCQPKEPPLTNAIKDALA' <MOR2>
A;Cross-references: EMBL:X93035; NID:g1085065; PIDN:CAA63603.1; PID:g1085066
A;Note: the differences at the carboxyl end are due to a frameshift error
C;Genetics:
A;Gene: brp39
C;Superfamily: Streptomyces chitinase chi40
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-405/Product: breast-regressing protein brp39 #status predicted <MAT>

  Query Match             74.4%;  Score 32;  DB 2;  Length 405;
  Best Local Similarity   100.0%;  Pred. No. 37;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTLASSE 7
              |||||||
Db        266 FTLASSE 272


RESULT 8
A84782
hypothetical protein At2g36560 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C;Accession: A84782
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: A84782
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-574 <STO>
A;Cross-references: GB:AE002093; NID:g4581154; PIDN:AAD24638.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g36560
A;Map position: 2

  Query Match             74.4%;  Score 32;  DB 2;  Length 574;
  Best Local Similarity   87.5%;  Pred. No. 54;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 TLASSETG 9
              |.||| :|
Db        290 TLASSSTG 297


RESULT 9
AB2209
two-component sensor histidine kinase alr3225 [imported] - Nostoc sp. (strain PCC 7
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C;Accession: AB2209
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iri
 Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Taba
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AB2209
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-986 <KUR>
A;Cross-references: GB:BA000019; PIDN:BAB74924.1; PID:g17132320; GSPDB:GN00179
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: alr3225

  Query Match             74.4%;  Score 32;  DB 2;  Length 986;
  Best Local Similarity   66.7%;  Pred. No. 99;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FTLASSETG 9
              ||: :||||
Db        536 FIVKNSETG 544


RESULT 10
S69542
hypothetical protein 34 - phage HP1
C;Species: phage HP1
C;Date: 06-Dec-1996 #sequence_revision 13-Mar-1997 #text_change 08-Oct-1999
C;Accession: S69542
R;Esposito, D.; Fitzmaurice, W.P.; Benjamin, R.C.; Goodman, S.D.; Waldman, A.S.; So
Nucleic Acids Res. 24, 2360-2368, 1996
A;Title: The complete nucleotide sequence of bacteriophage HP1 DNA.
A;Reference number: S69503; MUID:96279738; PMID:8710508
A;Accession: S69542
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-187 <ESP>
A;Cross-references: EMBL:U24159; NID:g1046235; PIDN:AAB09221.1; PID:g1046263
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, April 1995

  Query Match             72.1%;  Score 31;  DB 2;  Length 187;
  Best Local Similarity   66.7%;  Pred. No. 26;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 FTLASSETG 9
              | || :|||
Db         72 FNLAEAETG 80


RESULT 11
G87353
hypothetical protein CC0842 [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C;Accession: G87353
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelber
 B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; K
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser,
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: G87353
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-202 <STO>
A;Cross-references: GB:AE005673; NID:g13422095; PIDN:AAK22827.1; GSPDB:GN00148
C;Genetics:
A;Gene: CC0842
```

```
  Query Match             72.1%;  Score 31;  DB 2;  Length 202;
  Best Local Similarity   66.7%;  Pred. No. 28;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 FTLASSETG 9
              |:  | :|
Db        141 FTLVSGKTG 149


RESULT 12
T20178
hypothetical protein C53B4.6 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T20178
R;Berks, M.
submitted to the EMBL Data Library, December 1995
A;Reference number: Z19283
A;Accession: T20178
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-318 <WIL>
A;Cross-references: EMBL:Z68215; PIDN:CAA92454.1; GSPDB:GN00022; CESP:C53B4.6
A;Experimental source: clone C53B4
C;Genetics:
A;Gene: CESP:C53B4.6
A;Map position: 4
A;Introns: 25/2; 63/2; 97/3; 144/3; 200/3; 251/2

  Query Match             72.1%;  Score 31;  DB 2;  Length 318;
  Best Local Similarity   77.8%;  Pred. No. 47;
  Matches    7;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 FTLASSETG 9
              |||||  | |
Db        133 FTLASYEPG 141


RESULT 13
T24905
hypothetical protein T14C1.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C;Accession: T24905
R;Barlow, K.
submitted to the EMBL Data Library, April 1998
A;Reference number: Z19952
A;Accession: T24905
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-320 <WIL>
A;Cross-references: EMBL:Z70687; NID:g1256502; PIDN:CAA94616.1; GSPDB:GN00028; CESP:T14C
A;Experimental source: clone T14C1
C;Genetics:
A;Gene: CESP:T14C1.1
A;Map position: X
A;Introns: 75/3; 141/2; 189/1; 254/2

  Query Match             72.1%;  Score 31;  DB 2;  Length 320;
  Best Local Similarity   66.7%;  Pred. No. 47;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FTLASSETG 9
              |:|| ':||
Db         10 FSLACSDTG 18


RESULT 14
T20681
hypothetical protein F10A3.6 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T20681
R;Lloyd, C.
submitted to the EMBL Data Library, March 1997
A;Reference number: Z19309
A;Accession: T20681
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-324 <WIL>
A;Cross-references: EMBL:Z92829; PIDN:CAB07344.1; GSPDB:GN00023; CESP:F10A3.6
A;Experimental source: clone F10A3
C;Genetics:
A;Gene: CESP:F10A3.6
A;Map position: 5
A;Introns: 64/3; 94/1

  Query Match             72.1%;  Score 31;  DB 2;  Length 324;
  Best Local Similarity   75.0%;  Pred. No. 48;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 TLASSETG 9
              |: |||||
Db        235 TIGSSETG 242


RESULT 15
D83331
muconate cycloisomerase 1 PA2509 [imported] - Pseudomonas aeruginosa (strain PAO1)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C;Accession: D83331
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.;
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic p
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: D83331
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-373 <STO>
A;Cross-references: GB:AE004678; GB:AE004091; NID:g9948560; PIDN:AAG05897.1; GSPDB:G
A;Experimental source: strain PAO1
C;Genetics:
A;Gene: catB; PA2509
C;Superfamily: muconate cycloisomerase

  Query Match             72.1%;  Score 31;  DB 2;  Length 373;
  Best Local Similarity   66.7%;  Pred. No. 56;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FTLASSETG 9
              :|||| :||
Db        140 WTLASGDTG 148


Search completed: September 24, 2003, 13:49:18
Job time : 11.9213 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 24, 2003, 13:26:37 ; Search time 5.66292 Seconds
(without alignments)
74.739 Million cell updates/sec

Title: US-09-744-282-2
Perfect score: 43
Sequence: 1 PTLASSETG 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	43	100.0	383	1	C3L1_HUMAN	P36222 homo sapien
2	34	79.1	390	1	C3L2_HUMAN	Q15782 homo sapien
3	32	74.4	360	1	YACK_BACSU	P37573 bacillus su
4	32	74.4	381	1	C3L1_MOUSE	Q61362 mus musculu
5	31	72.1	187	1	YO34_BPHP1	P51738 bacteriopha
6	31	72.1	268	1	1NHA_MYCAV	O07400 mycobacteri
7	31	72.1	402	1	CN7B_SCHPO	Q09722 schizosacch
8	31	72.1	445	1	FLDL_VIBPA	Q03475 vibrio para
9	31	72.1	459	1	VL2_HPV6A	Q84297 human papil
10	31	72.1	459	1	VL2_HPV6B	P03106 human papil
11	31	72.1	590	1	VG29_BPT4	P13337 bacteriopha
12	30	69.8	246	1	PCL1_CHICK	P10399 gallus gall
13	30	69.8	317	1	MSHR_BOVIN	P47798 bos taurus
14	30	69.8	782	1	SM4B_MOUSE	Q62179 mus musculu
15	29	67.4	49	1	R334_ENTFA	P59629 enterococcu
16	29	67.4	139	1	STBB_PSESM	Q52562 pseudomonas
17	29	67.4	178	1	IDIX_RAT	P25086 rattus norv
18	29	67.4	318	1	YJTA_ECOLI	P24203 escherichia
19	29	67.4	342	1	YEA4_YEAST	P40064 saccharomyc
20	29	67.4	344	1	ASTA_ECOLI	P76218 escherichia
21	29	67.4	372	1	GPR_BACHD	Q9kd78 bacillus ha
22	29	67.4	401	1	YOPB_YERPN	P37131 yersinia en
23	29	67.4	401	1	YOPB_YERPS	Q06114 yersinia ps
24	29	67.4	486	1	EHS1_SCHPO	Q10063 schizosacch
25	29	67.4	573	1	RECN_TREPA	O83456 treponema p
26	29	67.4	599	1	VG9_BPPZA	P07534 bacteriopha
27	29	67.4	642	1	FLID_CAMJE	Q9phw6 campylobact
28	29	67.4	659	1	AMYA_STRRN	P18791 streptococc
29	29	67.4	920	1	YML1_YEAST	Q03758 saccharomyc
30	29	67.4	1076	1	NUP1_YEAST	P20676 saccharomyc
31	29	67.4	1642	1	KA43_HUMAN	Q96ay4 homo sapien
32	29	67.4	2388	1	SPCP_RAT	Q9qwn8 rattus norv
33	28	65.1	104	1	YBK0_YEAST	P38168 saccharomyc
34	28	65.1	113	1	YC71_MYCTU	Q11048 mycobacteri
35	28	65.1	129	1	Y4PG_RHISN	P55445 rhizobium s
36	28	65.1	137	1	YJXK_BACSU	P40761 bacillus su
37	28	65.1	218	1	YBA4_STRCO	Q9rkm3 streptomyce
38	28	65.1	234	1	TRAR_AGRVT	P33909 agrobacteri
39	28	65.1	285	1	CY1_RHOSH	Q02760 rhodobacter
40	28	65.1	304	1	UPK_CLOPE	Q8x156 clostridium
41	28	65.1	317	1	MSHR_CANFA	O77616 canis famil
42	28	65.1	317	1	MSHR_VULVU	Q29154 vulpes vulp
43	28	65.1	321	1	ISP_BACCS	P29140 bacillus cl
44	28	65.1	322	1	GVPL_HALME	Q02237 halobacteri
45	28	65.1	339	1	RSPB_ECOLI	P38105 escherichia

ALIGNMENTS

RESULT 1
C3L1_HUMAN
ID C3L1_HUMAN STANDARD; PRT; 383 AA.
AC P36222; P30923;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Chitinase-3 like protein 1 precursor (Cartilage glycoprotein-39) (GP-
DE 39) (39 kDa synovial protein) (YKL-40).
GN CHI3L1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Cartilage;
RX MEDLINE=94064658; PubMed=8245017;
RA Hakala B.E., White C., Recklies A.D.;
RT "Human cartilage gp-39, a major secretory product of articular
RT chondrocytes and synovial cells, is a mammalian member of a chitinase
RT protein family.";
RL J. Biol. Chem. 268:25803-25810(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=97386591; PubMed=9244440;
RA Rehli M., Krause S.W., Andreesen R.;
RT "Molecular characterization of the gene for human cartilage gp-39
RT (CHI3L1), a member of the chitinase protein family and marker for
RT late stages of macrophage differentiation.";
RL Genomics 43:221-225(1997).
RN [3]
RP SEQUENCE OF 22-45.
RX MEDLINE=90328983; PubMed=2375755;
RA Nyirkos P., Golds E.E.;
RT "Human synovial cells secrete a 39 kDa protein similar to a bovine
RT mammary protein expressed during the non-lactating period.";
RL Biochem. J. 269:265-268(1990).
CC -!- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN THE CAPACITY OF CELLS TO
CC RESPOND TO AND COPE WITH CHANGES IN THEIR ENVIRONMENT.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- TISSUE SPECIFICITY: PRESENT IN ARTICULAR CHONDROCYTES, SYNOVIAL
CC CELLS AS WELL AS IN LIVER. UNDETECTABLE IN MUSCLE TISSUES, LUNG,
CC PANCREAS, MONONUCLEAR CELLS, OR FIBROBLASTS.
CC -!- PTM: GLYCOSYLATED.
CC -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.
CC ----------

CC ----------

```
DR   EMBL; M80927; AAA16074.1; -.
DR   EMBL; Y08374; CAA69661.1; -.
DR   EMBL; Y08375; CAA69661.1; JOINED.
DR   EMBL; Y08376; CAA69661.1; JOINED.
DR   EMBL; Y08377; CAA69661.1; JOINED.
DR   EMBL; Y08378; CAA69661.1; JOINED.
DR   PIR; A49562; A49562.
DR   PDB; 1LA7; 10-APR-02.
DR   Genew; HGNC:1932; CHI3L1.
DR   MIM; 601525; -.
DR   GO; GO:0005578; C:extracellular matrix; TAS.
DR   GO; GO:0005615; C:extracellular space; TAS.
DR   GO; GO:0005207; F:extracellular matrix glycoprotein; TAS.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   InterPro; IPR001579; Glyco_hydro_18/2.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   ProDom; PD000471; Glyco_hydro_18; 1.
DR   SMART; SM00636; Glyco_18; 1.
DR   PROSITE; PS01095; CHITINASE_18; FALSE_NEG.
KW   Glycoprotein; Signal; 3D-structure.
FT   SIGNAL        1     21
FT   CHAIN        22    383       CHITINASE-3 LIKE PROTEIN 1.
FT   CARBOHYD     60     60       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   383 AA;  42613 MW;  75ADD8298EEEC2D1 CRC64;

  Query Match             100.0%;  Score 43;  DB 1;  Length 383;
  Best Local Similarity   100.0%;  Pred. No. 0.055;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTLASSETG 9
              |||||||||
Db        265 FTLASSETG 273


RESULT 2
C3L2_HUMAN
ID   C3L2_HUMAN     STANDARD;      PRT;   390 AA.
AC   Q15782; Q15749; Q15783;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Chitinase 3-like protein 2 precursor (YKL-39) (Chondrocyte protein
DE   39).
GN   CHI3L2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS LONG AND SHORT).
RA   Grossman A., Matsuyama T., Baker E., Waterhouse P., Sutherland G.R.,
RA   Mak T.W.;
RT   "Cloning of a novel lymphoid restricted human chitinase and
RT   localization to 1p13.3.";
RL   Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM SHORT).
RC   TISSUE=Articular cartilage;
RX   MEDLINE=96325055; PubMed=8702629;
RA   Hu B., Trinh K., Figueira W.F., Price P.A.;
RT   "Isolation and sequence of a novel human chondrocyte protein related
RT   to mammalian members of the chitinase protein family.";
RL   J. Biol. Chem. 271:19415-19420(1996).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=Short;
CC         IsoId=Q15782-1; Sequence=Displayed;
CC       Name=Long;
CC         IsoId=Q15782-2; Sequence=VSP_001066;
CC         Note=No experimental confirmation available;
CC   -!- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN CHONDROCYTES, FOLLOWED
CC       BY SYNOVIOCYTES, LUNG AND HEART. NOT DETECTED IN BRAIN, SPLEEN,
CC       PANCREAS, AND LIVER.
CC   -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; U58515; AAB04534.1; -.
DR   EMBL; U58514; AAB04533.1; -.
DR   EMBL; U49835; AAC50597.1; ALT_INIT.
DR   Genew; HGNC:1933; CHI3L2.
DR   MIM; 601526; -.
DR   GO; GO:0005615; C:extracellular space; TAS.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   InterPro; IPR001579; Glyco_hydro_18/2.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   ProDom; PD000471; Glyco_hydro_18; 1.
DR   SMART; SM00636; Glyco_18; 1.
DR   PROSITE; PS01095; CHITINASE_18; FALSE_NEG.
KW   Hydrolase; Signal; Alternative splicing.
FT   SIGNAL        1     26       POTENTIAL.
FT   CHAIN        27    390       CHITINASE 3-LIKE PROTEIN 2.
FT   CARBOHYD     35     35       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC      1     12       MGATTMDQKSLW -> SQQLWLGSPDEVWLYLVCEHTHIFT
FT                                AIIWDSRTRFGSTDFSFHPST (in isoform Long).
FT                                /FTId=VSP_001066.
SQ   SEQUENCE   390 AA;  43500 MW;  97886A2F3AA35677 CRC64;

  Query Match             79.1%;  Score 34;  DB 1;  Length 390;
  Best Local Similarity   87.5%;  Pred. No. 6.2;
  Matches    7;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTLASSET 8
              |||:|:||
Db        273 FTLASART 280


RESULT 3
YACK_BACSU
ID   YACK_BACSU     STANDARD;      PRT;   360 AA.
AC   P37573;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Hypothetical protein yacK.
GN   YACK.
OS   Bacillus subtilis.
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX   NCBI_TaxID=1423;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168;
RX   MEDLINE=96051385; PubMed=7584024;
RA   Ogasawara N., Nakai S., Yoshikawa H.;
RT   "Systematic sequencing of the 180 kilobase region of the Bacillus
RT   subtilis chromosome containing the replication origin.";
RL   DNA Res. 1:1-14(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168;
RX   MEDLINE=98044033; PubMed=9384377;
RA   Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA   Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA   Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA   Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA   Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA   Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA   Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA   Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA   Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
```

```
RA   Guiseppi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA   Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA   Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA   Kobayashi Y., Koetter P., Koningstein G., Krogh S., Kumano M.,
RA   Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA   Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA   Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA   Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA   Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA   Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA   Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA   Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA   Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA   Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA   Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA   Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA   Viari A., Wambutt R., Wedler E., Wedler H., Weitzenegger T.,
RA   Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA   Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RT   "The complete genome sequence of the Gram-positive bacterium Bacillus
RT   subtilis.";
RL   Nature 390:249-256(1997).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D26185; BAA05322.1; -.
DR   EMBL; Z99104; CAB11864.1; -.
DR   PIR; S66117; S66117.
DR   SubtiList; BG10150; yacK.
DR   InterPro; IPR003390; DUF147.
DR   InterPro; IPR000445; HhH.
DR   InterPro; IPR003583; HHH_1.
DR   Pfam; PF02457; DUF147; 1.
DR   Pfam; PF00633; HHH; 1.
DR   SMART; SM00278; HhH1; 1.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   360 AA;  40734 MW;  92994B4A9160026A9 CRC64;

  Query Match             74.4%;  Score 32;  DB 1;  Length 360;
  Best Local Similarity   75.0%;  Pred. No. 16;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 FLASSEIG 9
              |:::|||||
Db        101 TISSSETG 108


RESULT 4
C3L1_MOUSE
ID   C3L1_MOUSE     STANDARD;      PRT:   381 AA.
AC   Q61362;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Chitinase-3 like protein 1 precursor (Cartilage glycoprotein-39)
DE   (GP-39) (BRP39 protein).
GN   CHI3L1 OR BRP39.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FVB/N; TISSUE=Breast;
RX   MEDLINE=95060797; PubMed=7970700;
RA   Morrison B.W., Leder P.;
RT   "neu and ras initiate murine mammary tumors that share genetic markers
RT   generally absent in c-myc and int-2-initiated tumors.";
RL   Oncogene 9:3417-3426(1994).
CC   -!- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN THE CAPACITY OF CELLS TO
CC       RESPOND TO AND COPE WITH CHANGES IN THEIR ENVIRONMENT (BY
CC       SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Extracellular (By similarity).
CC   -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X93035; CAA63603.1; -.
DR   PIR; S61551; S61551.
DR   MGD; MGI:1340899; Chi3l1.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   InterPro; IPR001579; Glyco_hydro_18/2.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   ProDom; PD000471; Glyco_hydro_18; 1.
DR   SMART; SM00636; Glyco_18; 1.
DR   PROSITE; PS01095; CHITINASE_18; FALSE_NEG.
KW   Glycoprotein; Signal.
FT   SIGNAL        1     21       POTENTIAL.
FT   CHAIN        22    381       CHITINASE-3 LIKE PROTEIN 1.
FT   CARBOHYD     60     60       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   381 AA;  43001 MW;  BF6581E8184F0450 CRC64;

  Query Match             74.4%;  Score 32;  DB 1;  Length 381;
  Best Local Similarity   100.0%;  Pred. No. 17;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTLASSE 7
              |||||||
Db        266 FTLASSE 272


RESULT 5
YO34_BPHP1
ID   YO34_BPHP1     STANDARD;      PRT:   187 AA.
AC   P51738;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical 20.8 kDa protein in LYS 3'region (ORF34).
OS   Bacteriophage HP1.
OC   Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae.
OX   NCBI_TaxID=10690;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HP1C1;
RX   MEDLINE=96279738; PubMed=8710508;
RA   Esposito D., Fitzmaurice W.P., Benjamin R.C., Goodman S.D.,
RA   Waldman A.S., Scocca J.J.;
RT   "The complete nucleotide sequence of bacteriophage HP1 DNA.";
RL   Nucleic Acids Res. 24:2360-2368(1996).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U24159; AAB09221.1; -.
DR   PIR; S69542; S69542.
KW   Hypothetical protein.
SQ   SEQUENCE   187 AA;  20761 MW;  A9A073D71DDEC02E CRC64;
```

```
  Query Match           72.1%;  Score 31;  DB 1;  Length 187;
  Best Local Similarity  66.7%;  Pred. No. 13;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 FTLASSETG 9
              | ||: :|||
Db         72 FNLAEAETG 80


RESULT 6
INHA_MYCAV
ID   INHA_MYCAV     STANDARD;      PRT:   268 AA.
AC   O07400;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Enoyl-[acyl-carrier-protein] reductase [NADH] (EC 1.3.1.9) (NADH-
DE   dependent enoyl-ACP reductase).
GN   INHA.
OS   Mycobacterium avium.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX   NCBI_TaxID=1764;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=GIR10;
RX   MEDLINE=98195739; PubMed=9534249;
RA   Labo M., Gusberti L., de Rossi E., Speziale P., Riccardi G.;
RT   "Determination of a 15437 bp nucleotide sequence around the inhA gene
RT   of Mycobacterium avium and similarity analysis of the products of
RT   putative ORFs.";
RL   Microbiology 144:807-814(1998).
CC   -!- CATALYTIC ACTIVITY: Acyl-[acyl-carrier protein] + NAD(+) = trans-
CC       2,3-dehydroacyl-[acyl-carrier protein] + NADH.
CC   -!- PATHWAY: SECOND REDUCTIVE STEP IN FATTY ACID BIOSYNTHESIS. THIS
CC       ISOZYME IS INVOLVED IN MYCOLIC ACID BIOSYNTHESIS.
CC   -!- SUBUNIT: Homotetramer (By similarity).
CC   -!- SIMILARITY: SOME SIMILARITY TO THE SHORT-CHAIN
CC       DEHYDROGENASES/REDUCTASES (SDR) FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AF002133; AAC46204.1; -.
DR   HSSP; P46533; 1ENY.
DR   InterPro; IPR002198; ADH_short.
DR   Pfam; PF00106; adh_short; 1.
KW   Oxidoreductase; NAD; Fatty acid biosynthesis.
FT   NP_BIND     135    164       NAD (POTENTIAL).
SQ   SEQUENCE   268 AA;  28529 MW;  F73501BD2B0F9990 CRC64;

  Query Match           72.1%;  Score 31;  DB 1;  Length 268;
  Best Local Similarity  75.0%;  Pred. No. 19;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 TLASSETG 9
              ::|||.|;
Db         94 SIASSETG 101


RESULT 7
CN7B_SCHPO
ID   CN7B_SCHPO     STANDARD;      PRT:   402 AA.
AC   Q09722; Q9US29;
DT   01-NOV-1995 (Rel. 32, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   COP9/signalosome complex subunit 7B.
GN   CSN7B OR SPAC1751.03 OR SPAC31A2.01.
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
OC   Schizosaccharomyces.
OX   NCBI_TaxID=4896;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972;
RX   MEDLINE=21848401; PubMed=11859360;
RA   Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA   Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA   Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA   Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA   Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA   Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA   James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA   Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA   Oliver K., O'Neil S., Pearson D., Quail M.A., Rabbinowitsch E.,
RA   Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA   Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA   Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA   Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA   Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA   Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,
RA   Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA   Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA   Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA   Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA   Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA   Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA   Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA   Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA   Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT   "The genome sequence of Schizosaccharomyces pombe.";
RL   Nature 415:871-880(2002).
CC   -!- SIMILARITY: Contains 1 PCI domain.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AL133154; CAB61449.1; -.
DR   EMBL; Z50113; CAA90459.1; -.
DR   GeneDB_SPombe; SPAC1751.03; -.
DR   InterPro; IPR000717; PCI.
DR   Pfam; PF01399; PCI; 1.
DR   SMART; SM00088; PINT; 1.
KW   Signalosome.
FT   DOMAIN      274    365       PCI.
SQ   SEQUENCE   402 AA;  45085 MW;  CD972E781BFC0A3C CRC64;

  Query Match           72.1%;  Score 31;  DB 1;  Length 402;
  Best Local Similarity  66.7%;  Pred. No. 31;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FTLASSETG 9
              ||:|||: |
Db        383 FTIASSKKG 391


RESULT 8
FLDL_VIBPA
ID   FLDL_VIBPA     STANDARD;      PRT:   445 AA.
AC   Q03475;
DT   01-OCT-1993 (Rel. 27, Created)
DT   15-SEP-2003 (Rel. 42, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Lateral flagellar hook-associated protein 2 (HAP2) (Filament CAP
```

```
DE   protein) (Flagellar cap protein).
GN   FLIDL OR LAFB OR VPA1550.
OS   Vibrio parahaemolyticus.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC   Vibrionaceae; Vibrio.
OX   NCBI_TaxID=670;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BB22;
RX   MEDLINE=93273702; PubMed-8501040;
RA   McCarter L.L., Wright M.E.;
RT   "Identification of genes encoding components of the swarmer cell
RT   flagellar motor and propeller and a sigma factor controlling
RT   differentiation of Vibrio parahaemolyticus.";
RL   J. Bacteriol. 175:3361-3371(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=RIMD 2210633 / Serotype O3:K6;
RX   MEDLINE-22508454; PubMed-12620739;
RA   Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
RA   Iijima Y., Najima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,
RA   Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;
RT   "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
RT   distinct from that of V. cholerae.";
RL   Lancet 361:743-749(2003).
CC   -!- FUNCTION: REQUIRED FOR THE MORPHOGENESIS AND FOR THE ELONGATION OF
CC       THE FLAGELLAR FILAMENT BY FACILITATING POLYMERIZATION OF THE
CC       FLAGELLIN MONOMERS AT THE TIP OF GROWING FILAMENT. FORMS A CAPPING
CC       STRUCTURE, WHICH PREVENTS FLAGELLIN SUBUNITS (TRANSPORTED THROUGH
CC       THE CENTRAL CHANNEL OF THE FLAGELLUM) FROM LEAKING OUT WITHOUT
CC       POLYMERIZATION AT THE DISTAL END. ESSENTIAL FOR SWARMING MOTILITY.
CC   -!- SUBUNIT: Homopentamer (By similarity).
CC   -!- SUBCELLULAR LOCATION: Flagellar.
CC   -!- MISCELLANEOUS: V.PARAHAEMOLYTICUS POSSESSES TWO FLAGELLAR SYSTEMS:
CC       A SINGLE POLAR FLAGELLUM PROPELS THE BACTERIUM IN LIQUID
CC       (SWIMMING), WHILE MULTIPLE LATERAL (PERITRICHOUS) FLAGELLA MOVE
CC       THE BACTERIUM OVER SURFACES (SWARMING). THE POLAR FLAGELLUM IS
CC       SYNTHESIZED CONSTITUTIVELY BUT LATERAL FLAGELLA ARE PRODUCED ONLY
CC       UNDER CONDITIONS IN WHICH THE POLAR FLAGELLUM IS NOT FUNCTIONAL.
CC   -!- SIMILARITY: BELONGS TO THE FLID FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L06176; AAA27530.1; -.
DR   EMBL; U52957; AAB07351.1; -.
DR   EMBL; AP005089; BAC62893.1; -.
DR   PIR; B40590; B40590.
DR   InterPro; IPR003481; FliD.
DR   Pfam; PF02465; FliD; 1.
KW   Flagella; Coiled coil; Complete proteome.
FT   DOMAIN      388    423       COILED COIL (POTENTIAL).
FT   CONFLICT    226    226       K -> N (IN REF. 1).
SQ   SEQUENCE   445 AA;  48279 MW;  7B1836FC8722F90A CRC64;

  Query Match           72.1%;  Score 31;  DB 1;  Length 445;
  Best Local Similarity  66.7%;  Pred. No. 34;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FILASSETG 9
              | |:|:|||
Db        177 FMLSSIETG 185


RESULT 9
VL2_HPV6A
ID   VL2_HPV6A      STANDARD;      PRT:   459 AA.
AC   Q84297;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   Minor capsid protein L2.
GN   L2.
OS   Human papillomavirus type 6a.
OC   Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC   Papillomavirus.
OX   NCBI_TaxID=37122;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95297152; PubMed-7778283;
RA   Hofmann K.J., Cook J.C., Joyce J.G., Brown D.R., Schultz L.D.,
RA   George H.A., Rosolowsky M., Fife K.H., Jansen K.U.;
RT   "Sequence determination of human papillomavirus type 6a and assembly
RT   of virus-like particles in Saccharomyces cerevisiae.";
RL   Virology 209:506-518(1995).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L41216; AAA74217.1; -.
DR   InterPro; IPR000784; Late_L2.
DR   Pfam; PF00513; late_protein_L2; 1.
KW   Coat protein; Late protein.
SQ   SEQUENCE   459 AA;  49406 MW;  FA1523CACBC89406 CRC64;

  Query Match           72.1%;  Score 31;  DB 1;  Length 459;
  Best Local Similarity  75.0%;  Pred. No. 36;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FILASSET 8
              |:: ||.:
Db        129 FTITSSET 136


RESULT 10
VL2_HPV6B
ID   VL2_HPV6B      STANDARD;      PRT:   459 AA.
AC   P03106;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   Minor capsid protein L2.
GN   L2.
OS   Human papillomavirus type 6b.
OC   Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC   Papillomavirus.
OX   NCBI_TaxID=10600;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=84131949; PubMed=6321162;
RA   Schwarz E., Durst M., Demankowski C., Lattermann O., Zech R.,
RA   Wolfsperger E., Suhai S., Zur Hausen H.;
RT   "DNA sequence and genome organization of genital human papillomavirus
RT   type 6b.";
RL   EMBO J. 2:2341-2348(1983).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X00203; CAA25025.1; -.
DR   PIR; A03647; P2WL6.
```

```
DR   InterPro; IPR000784; Late_L2.
DR   Pfam; PF00513; late_protein_L2; 1.
KW   Coat protein; Late protein.
SQ   SEQUENCE   459 AA;  49391 MW;  15048EBC90C48E46 CRC64;

  Query Match           72.1%;  Score 31;  DB 1;  Length 459;
  Best Local Similarity   75.0%;  Pred. No. 36;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FTLASSET 8
              || : ||| |
Db        129 FTIISSET 136


RESULT 11
VG29_BPT4
ID   VG29_BPT4      STANDARD;      PRT;   590 AA.
AC   P13337;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-JAN-1990 (Rel. 13, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Tail-tube assembly protein Gp29 (Tail length regulator)
DE   (Folylpolyglutamate synthase) (EC 6.3.2.17).
GN   29.
OS   Bacteriophage T4.
OC   Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;
OC   T4-like viruses.
OX   NCBI_TaxID=10665;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88206086; PubMed=3363870;
RA   Ishimoto L.K., Ishimoto K.S., Cascino A., Cipollaro M.,
RA   Eiserling F.A.;
RT   "The structure of three bacteriophage T4 genes required for tail-tube
RT   assembly.";
RL   Virology 164:81-90(1988).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Kutter E., Arisaka F., Kunisawa T., Tsugita A., Mosig G.,
RA   Mesyanzhinov V., Ruger W., Stidham T., Thomas E.;
RT   "Bacteriophage T4 genome analysis.";
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: STRUCTURAL COMPONENT OF THE BASEPLATE OF THE PHAGE AND
CC       FUNCTIONS AS A STRUCTURAL INITIATOR OF THE CENTRAL HUB ASSEMBLY OF
CC       THE BASEPLATE. IT ALSO CATALYZES FOLYL POLYGLUTAMATE SYNTHETASE
CC       ACTIVITY.
CC   -!- CATALYTIC ACTIVITY: ATP + (tetrahydrofolyl-[Glu])(N) + L-glutamate
CC       = ADP + phosphate + (tetrahydrofolyl-[Glu])(N+1).
CC   ---------------------------------------- ------------ ------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------- ------------ -- ------------
DR   EMBL; M20298; AAA32538.1; -.
DR   EMBL; AF158101; AAD42437.1; -.
DR   PIR; JF0035; GBBPT4.
KW   Structural protein; Late protein; Ligase.
SQ   SEQUENCE   590 AA;  64410 MW;  F5E6161019E6F9B4 CRC64;

  Query Match           72.1%;  Score 31;  DB 1;  Length 590;
  Best Local Similarity   77.8%;  Pred. No. 46;
  Matches    7;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 FTLASSETG 9
              ||  |||||
Db        273 FTKFSSETG 281


RESULT 12
POL1_CHICK
ID   POL1_CHICK     STANDARD;      PRT;   246 AA.
AC   P10399;
DT   01-MAR-1989 (Rel. 10, Created)
DT   01-MAR-1989 (Rel. 10, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Retrovirus-related POL polyprotein [Contains: Reverse transcriptase
DE   (EC 2.7.7.49); Endonuclease] (Fragments).
GN   POL.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=85270471; PubMed=2410912;
RA   Dunwiddie C., Faras A.J.;
RT   "Presence of retrovirus reverse transcriptase-related gene sequences
RT   in avian cells lacking endogenous avian leukosis viruses.";
RL   Proc. Natl. Acad. Sci. U.S.A. 82:5097-5101(1985).
RN   [2]
RP   REVISION TO 2.
RA   Dunwiddie C., Faras A.J.;
RL   Submitted (OCT-1985) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC       + (DNA)(N).
CC   ---------------------------------------- ------------ ------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------- ------------ ------------------
DR   EMBL; K03271; -; NOT_ANNOTATED_CDS.
DR   EMBL; K03272; -; NOT_ANNOTATED_CDS.
DR   PIR; A38044; A25779.
DR   PDB; 1M0C; 26-JUN-02.
KW   Hydrolase; Transferase; RNA-directed DNA polymerase; Nuclease;
KW   Endonuclease; Polyprotein; 3D-structure.
FT   NON_TER       1      1
FT   NON_CONS    111    112
FT   NON_TER     246    246
SQ   SEQUENCE   246 AA;  26354 MW;  EA036F7C0792C7C6 CRC64;

  Query Match           69.8%;  Score 30;  DB 1;  Length 246;
  Best Local Similarity   77.8%;  Pred. No. 29;
  Matches    7;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 FTLASSETG 9
              || ||| ||
Db         98 FTDASSSTG 106


RESULT 13
MSHR_BOVIN
ID   MSHR_BOVIN     STANDARD;      PRT;   317 AA.
AC   P47798; Q28025;
DT   01-FEB-1996 (Rel. 33, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Melanocyte stimulating hormone receptor (MSH-R) (Melanotropin
DE   receptor) (Melanocortin-1 receptor) (MC1-R) (BDF3).
GN   MC1R OR MSHR.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
```

```
RX   MEDLINE=94307438; PubMed=8034052;
RA   Vanetti M., Schoenrock C., Meyerhof W., Hoellt V.;
RT   "Molecular cloning of a bovine MSH receptor which is highly expressed
RT   in the testis.";
RL   FEBS Lett. 348:268-272(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Holstein;
RX   MEDLINE=96325106; PubMed=8661706;
RA   Joerg H., Fries H.R., Meijerink E., Stranzinger G.F.;
RT   "Red coat color in Holstein cattle is associated with a deletion in
RT   the MSHR gene.";
RL   Mamm. Genome 7:317-318(1996).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20093783; PubMed=10629296;
RA   Klungland H., Rood K.H., Nesbo C.L., Jakobsen K.S., Vage D.I.;
RT   "The melanocyte-stimulating hormone receptor (MC1-R) gene as a tool
RT   in evolutionary studies of artiodactyles.";
RL   Hereditas 131:39-45(1999).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Holstein;
RA   Kriegesmann B., Brenig B., Dierkes B., Jansen S.;
RT   "A new melanocortin 1-receptor allele is coupled with Agouti coat
RT   colour pattern in Holstein cattle.";
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: RECEPTOR FOR MSH (ALPHA, BETA) AND ACTH. DOES NOT SEEM
CC       TO BE ACTIVE WITH GAMMA-MSH. THE ACTIVITY OF THIS RECEPTOR IS
CC       MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYLATE CYCLASE. COULD BE
CC       INVOLVED IN SPERMATOGENESIS.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN THE TESTIS.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; S71017; AAB31361.1; -.
DR   EMBL; U39469; AAC48590.2; -.
DR   EMBL; Y13957; CAA74291.1; -.
DR   EMBL; Y19103; CAB64818.1; -.
DR   PIR; S45708; S45708.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Phosphorylation; Lipoprotein; Palmitate.
FT   DOMAIN        1     37       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     38     63       1 (POTENTIAL).
FT   DOMAIN       64     72       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     73     93       2 (POTENTIAL).
FT   DOMAIN       94    118       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    119    140       3 (POTENTIAL).
FT   DOMAIN      141    163       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    164    183       4 (POTENTIAL).
FT   DOMAIN      184    191       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    192    211       5 (POTENTIAL).
FT   DOMAIN      212    240       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    241    266       6 (POTENTIAL).
FT   DOMAIN      267    279       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    280    300       7 (POTENTIAL).
FT   DOMAIN      301    317       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD     29     29       N-LINKED (GLCNAC...) (POTENTIAL).
FT   LIPID       315    315       PALMITATE (POTENTIAL).
FT   CONFLICT     40     41       IP -> SL (IN REF. 1).
FT   CONFLICT     99     99       L -> P (IN REF. 1 AND 2).
SQ   SEQUENCE   317 AA;  34916 MW;  123692614F7FF168 CRC64;

  Query Match             69.8%;  Score 30;  DB 1;  Length 317;
  Best Local Similarity   66.7%;  Pred. No. 39;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 FTLASSETG 9
              |||| : ||
Db         24 FTLAPNRTG 32


RESULT 14
SM4B_MOUSE
ID   SM4B_MOUSE     STANDARD;      PRT;   782 AA.
AC   Q62179;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Semaphorin 4B (Semaphorin C) (Sema C) (Fragment).
GN   SEMA4B OR SEMAC OR SEMC.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NMRI; TISSUE=Brain;
RX   MEDLINE=95267431; PubMed=7748561;
RA   Pueschel A.W., Adams R.H., Betz H.;
RT   "Murine semaphorin D/collapsin is a member of a diverse gene family
RT   and creates domains inhibitory for axonal extension.";
RL   Neuron 14:941-948(1995).
RN   [2]
RP   INTERACTION WITH GIPC.
RX   MEDLINE=99253973; PubMed=10318831;
RA   Wang L.-H., Kalb R.G., Strittmatter S.M.;
RT   "A PDZ protein regulates the distribution of the transmembrane
RT   semaphorin, M-SemF.";
RL   J. Biol. Chem. 274:14137-14146(1999).
CC   -!- FUNCTION: INHIBITS AXONAL EXTENSION BY PROVIDING LOCAL SIGNALS TO
CC       SPECIFY TERRITORIES INACCESSIBLE FOR GROWING AXONS.
CC   -!- SUBUNIT: INTERACTS WITH GIPC PDZ DOMAIN.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- DEVELOPMENTAL STAGE: EXPRESSED FROM DAY 10 IN THE EMBRYO. LOW
CC       LEVELS FOUND BETWEEN DAYS 10-12. EXPRESSION PEAKS ON DAY 13 WITH
CC       MODERATE LEVELS FROM THEN UNTIL BIRTH.
CC   -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC   -!- SIMILARITY: Contains 1 Sema domain.
CC   -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X85992; CAA59984.1; -.
DR   PIR; I48746; I48746.
DR   MGD; MGI:107559; Sema4b.
DR   InterPro; IPR003659; Plexin-like.
DR   InterPro; IPR002165; Plexin_repeat.
DR   InterPro; IPR001627; Sema.
DR   Pfam; PF01437; PSI; 1.
DR   Pfam; PF01403; Sema; 1.
DR   SMART; SM00423; PSI; 1.
DR   SMART; SM00630; Sema; 1.
KW   Transmembrane; Immunoglobulin domain; Multigene family; Neurogenesis;
KW   Developmental protein; Glycoprotein.
FT   NON_TER       1      1
FT   DOMAIN       <1    662       EXTRACELLULAR (POTENTIAL).
```

```
FT   TRANSMEM    663    683       POTENTIAL.
FT   DOMAIN      684    782       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       16    483       SEMA.
FT   DOMAIN      471    525       PSI.
FT   DOMAIN      546    606       IG-LIKE C2-TYPE.
FT   DOMAIN      703    726       PRO-RICH.
FT   DISULFID    555    601       BY SIMILARITY.
FT   CARBOHYD     12     12       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     15     15       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     42     42       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     88     88       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    356    356       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    471    471       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    526    526       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    574    574       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    639    639       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   782 AA;  86823 MW;  627A81FC8F8F7AC8 CRC64;

  Query Match             69.8%;  Score 30;  DB 1;  Length 782;
  Best Local Similarity   66.7%;  Pred. No. 1.1e+02;
  Matches    6;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 FTLASSETG 9
              ||||  | :
Db        112 FTLAQDEAG 120


RESULT 15
R334_ENTFA
ID   R334_ENTFA     STANDARD;      PRT:    49 AA.
AC   P59629;
DT   15-SEP-2003 (Rel. 42, Created)
DT   15-SEP-2003 (Rel. 42, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   50S ribosomal protein L33 type 4.
GN   RPMG4 OR RPMG-4 OR EF3203.
OS   Enterococcus faecalis (Streptococcus faecalis).
OC   Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
OX   NCBI_TaxID=1351;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=V583 / ATCC 700802;
RX   MEDLINE=22550857; PubMed=12663927;
RA   Paulsen I.T., Banerjei L., Myers G.S.A., Nelson K.E., Seshadri R.,
RA   Read T.D., Fouts D.E., Eisen J.A., Gill S.R., Heidelberg J.F.,
RA   Tettelin H., Dodson R.J., Umayam L., Brinkac L., Beanan M.,
RA   Daugherty S., DeBoy R.T., Durkin S., Kolonay J., Madupu R., Nelson W.,
RA   Vamathevan J., Tran B., Upton J., Hansen T., Shetty J., Khouri H.,
RA   Utterback T., Radune D., Ketchum K.A., Dougherty B.A., Fraser C.M.;
RT   "Role of mobile DNA in the evolution of vancomycin-resistant
RT   Enterococcus faecalis.";
RL   Science 299:2071-2074(2003).
CC   -!- SIMILARITY: BELONGS TO THE L33P FAMILY OF RIBOSOMAL PROTEINS.
CC   ----------------------------------------- ----------------- - - - -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------- ----------------------------------
DR   EMBL; AE016957; AAO82875.1; -.
DR   TIGR; EF3203; -.
DR   HAMAP; MF_00294; -; 1.
DR   PROSITE; PS00582; RIBOSOMAL_L33; FALSE_NEG.
KW   Ribosomal protein.
SQ   SEQUENCE   49 AA;  5811 MW;  866C53A8E5A83145 CRC64;

  Query Match             67.4%;  Score 29;  DB 1;  Length 49;
  Best Local Similarity   75.0%;  Pred. No. 7.4;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 TLASSETG 9
              ||| :|||
Db          6 TLACAETG 13


Search completed: September 24, 2003, 13:42:24
Job time : 6.66292 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 24, 2003, 13:33:45 ; Search time 29.2247 Seconds
(without alignments)
79.469 Million cell updates/sec

Title: US-09-744-282-2
Perfect score: 43
Sequence: 1 FTLASSETG 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	43	100.0	383	4	Q96HI7	Q96hi7 homo sapien
2	43	100.0	383	4	Q8IVA4	Q8iva4 homo sapien
3	37	86.0	441	2	O52483	O52483 serratia ma
4	35	81.4	481	2	O05619	O05619 pseudomonas
5	34	79.1	390	4	Q96F97	Q96f97 homo sapien
6	34	79.1	931	5	Q8IS90	Q8is90 leishmania
7	33	76.7	224	17	Q8THV9	Q8thv9 methanosarc
8	33	76.7	383	6	Q8SPQ0	Q8spq0 capra hircu
9	33	76.7	383	6	Q29411	Q29411 sus scrofa
10	33	76.7	1265	4	Q9P2G7	Q9p2g7 homo sapien
11	33	76.7	2162	4	Q8N3U0	Q8n3u0 homo sapien
12	32	74.4	251	17	O27403	O27403 methanobact
13	32	74.4	261	17	Q9HPA0	Q9hpa0 halobacteri
14	32	74.4	352	11	Q9WTV1	Q9wtv1 rattus norv
15	32	74.4	372	13	Q90Y50	Q90y50 brachydanio
16	32	74.4	381	11	Q99J84	Q99j84 mus musculu
17	32	74.4	389	11	Q8BKL8	Q8bkl8 mus musculu
18	32	74.4	446	10	Q9SXN2	Q9sxn2 pyrus pyrif
19	32	74.4	565	16	Q8XJE4	Q8xje4 clostridium
20	32	74.4	574	10	Q9SJQ2	Q9sjq2 arabidopsis
21	32	74.4	986	16	Q8YS66	Q8ys66 anabaena sp
22	32	74.4	1243	5	Q9NGT8	Q9ngt8 leishmania
23	31	72.1	186	9	Q94MX7	Q94mx7 haemophilus
24	31	72.1	202	5	Q9U8Z4	Q9u8z4 tachypleus
25	31	72.1	202	16	Q9A9W9	Q9a9w9 caulobacter
26	31	72.1	314	4	Q9UDQ5	Q9udq5 homo sapien
27	31	72.1	318	5	Q18800	Q18800 caenorhabdi
28	31	72.1	324	5	O45339	O45339 caenorhabdi
29	31	72.1	356	10	Q8LH32	Q8lh32 oryza sativ
30	31	72.1	373	2	Q93SR2	Q93sr2 pseudomonas
31	31	72.1	373	16	Q9I0X3	Q9i0x3 pseudomonas
32	31	72.1	395	2	Q8GAZ7	Q8gaz7 burkholderi
33	31	72.1	414	3	Q9URK6	Q9urk6 schizosacch
34	31	72.1	420	5	Q22485	Q22485 caenorhabdi
35	31	72.1	459	12	Q99A89	Q99a89 human papil
36	31	72.1	459	12	Q99A90	Q99a90 human papil
37	31	72.1	459	12	Q9QLN9	Q9qln9 human papil
38	31	72.1	491	5	Q9XUF6	Q9xuf6 caenorhabdi
39	31	72.1	515	2	Q9ZBD4	Q9zbd4 rhodococcus
40	31	72.1	559	3	Q8NKC4	Q8nkc4 penicillium
41	31	72.1	604	11	Q85YT2	Q85yt2 mus musculu
42	31	72.1	605	17	O26511	O26511 methanobact
43	31	72.1	637	10	Q9M7C9	Q9m7c9 abies grand
44	31	72.1	817	12	Q9J5H1	Q9j5h1 fowlpox vir
45	31	72.1	817	12	Q90761	Q90761 fowlpox vir

ALIGNMENTS

RESULT 1
Q96HI7
ID Q96HI7 PRELIMINARY; PRT; 383 AA.
AC Q96HI7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Similar to chitinase 3-like 1 (Cartilage glycoprotein-39).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC008568; AAH08568.1; -.
DR InterPro; IPR001223; Glyco_hydro_18.
DR InterPro; IPR005829; Sug_transporter.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR SMART; SM00636; Glyco_18; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
SQ SEQUENCE 383 AA; 42625 MW; 8C9EF133BDC7C2D1 CRC64;

Query Match 100.0%; Score 43; DB 4; Length 383;
Best Local Similarity 100.0%; Pred. No. 0.41;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 FTLASSETG 9
              |||||||||
Db        265 FTLASSETG 273
```

RESULT 2
Q8IVA4
ID Q8IVA4 PRELIMINARY; PRT; 383 AA.
AC Q8IVA4;

```
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Similar to chitinase 3-like 1 (Cartilage glycoprotein-39).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RA   Strausberg R.;
RL   Submitted (OCT-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC038354; AAH38354.1; -.
SQ   SEQUENCE   383 AA;  42526 MW;  76345433BDC7C6C0 CRC64;

  Query Match             100.0%;  Score 43;  DB 4;  Length 383;
  Best Local Similarity   100.0%;  Pred. No. 0.41;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTLASSETG 9
              |||||||||
Db        265 FTLASSETG 273


RESULT 3
O52483
ID   O52483      PRELIMINARY;      PRT;   441 AA.
AC   O52483;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Putative ATP-binding protein.
GN   WZT.
OS   Serratia marcescens.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Serratia.
OX   NCBI_TaxID=615;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=N28b;
RX   MEDLINE=99173913; PubMed=10074093;
RA   Saigi F., Climent N., Pique N., Sanchez C., Merino S., Rubires X.,
RA   Aguilar A., Tomas J.M., Regue M.;
RT   "Genetic analysis of the Serratia marcescens N28b O4 antigen gene
RT   cluster.";
RL   J. Bacteriol. 181:1883-1891(1999).
DR   EMBL; AF038816; AAC00182.1; -.
DR   InterPro; IPR003593; AAA_ATPase.
DR   InterPro; IPR003439; ABC_transporter.
DR   Pfam; PF00005; ABC_tran; 1.
DR   ProDom; PD000006; ABC_transporter; 1.
DR   SMART; SM00382; AAA; 1.
KW   ATP-binding.
SQ   SEQUENCE   441 AA;  48197 MW;  56F7B3F1BDC2AB5F CRC64;

  Query Match             86.0%;  Score 37;  DB 2;  Length 441;
  Best Local Similarity   100.0%;  Pred. No. 11;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTLASSET 8
              |||||||:
Db        332 FTLASSET 339


RESULT 4
O05619
ID   O05619      PRELIMINARY;      PRT;   481 AA.
AC   O05619;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Vanillin dehydrogenase.
GN   VDH.
OS   Pseudomonas sp.
OC   Bacteria; Proteobacteria.
OX   NCBI_TaxID=306;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HR199;
RX   MEDLINE=97252489; PubMed=9098058;
RA   Priefert H., Rabenhorst J., Steinbuechel A.;
RT   "Molecular characterization of genes of Pseudomonas sp. strain HR199
RT   involved in bioconversion of vanillin to protocatechuate.";
RL   J. Bacteriol. 179:2595-2607(1997).
CC   -!- SIMILARITY: BELONGS TO THE ALDEHYDE DEHYDROGENASES FAMILY.
DR   EMBL; Y11520; CAA72286.1; -.
DR   HSSP; P51977; 1BXS.
DR   InterPro; IPR002086; Aldehyde_dehydr.
DR   Pfam; PF00171; aldedh; 1.
DR   PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; 1.
KW   Oxidoreductase.
SQ   SEQUENCE   481 AA;  50836 MW;  9661EF2DF0F6D59E CRC64;

  Query Match             81.4%;  Score 35;  DB 2;  Length 481;
  Best Local Similarity   77.8%;  Pred. No. 33;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FTLASSETG 9
              || |:||||
Db         83 FIAAASETG 91


RESULT 5
Q96F97
ID   Q96F97      PRELIMINARY;      PRT;   390 AA.
AC   Q96F97;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Chitinase 3-like 2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RA   Strausberg R.;
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC011460; AAH11460.1; -.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   ProDom; PD000471; Glyco_hydro_18; 1.
DR   SMART; SM00636; Glyco_18; 1.
SQ   SEQUENCE   390 AA;  43529 MW;  6801E4AF721AD63F CRC64;

  Query Match             79.1%;  Score 34;  DB 4;  Length 390;
  Best Local Similarity   87.5%;  Pred. No. 44;
  Matches    7;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTLASSET 8
              ||||||:|
Db        273 FTLASAET 280


RESULT 6
Q8IS90
ID   Q8IS90      PRELIMINARY;      PRT;   931 AA.
AC   Q8IS90;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   DNA polymerase I type A.
GN   POLIA.
OS   Leishmania major.
```

```
OC   Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX   NCBI_TaxID=5664;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Misra S., Chaudhuri G.;
RT   "Cloning of a PolIA-like mitochondrial DNA polymerase I gene (LmPOLIA)
RT   from Leishmania major.";
RL   Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY147013; AAN61165.1; -.
SQ   SEQUENCE   931 AA;  99947 MW;  B4E02B4FFC2949A4 CRC64;

  Query Match           79.1%;  Score 34;  DB 5;  Length 931;
  Best Local Similarity  77.8%;  Pred. No. 1.2e+02;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FTLASSETG 9
              |:|:| | |
Db         68 FSLASSEAG 76


RESULT 7
Q8THV9
ID   Q8THV9      PRELIMINARY;      PRT;   224 AA.
AC   Q8THV9;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Undecaprenyl diphosphate synthase.
GN   MA4402.
OS   Methanosarcina acetivorans.
OC   Archaea; Euryarchaeota; Methanococci; Methanosarcinales;
OC   Methanosarcinaceae; Methanosarcina.
OX   NCBI_TaxID=2214;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C2A / ATCC 35395 / DSM 2834;
RX   MEDLINE=21929760; PubMed=11932238;
RA   Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,
RA   FitzHugh W., Calvo S., Engels R., Smirnov S., Atnoor D., Brown A.,
RA   Allen N., Naylor J., Stange-Thomann N., DeArellano K., Johnson R.,
RA   Linton L., McEwan P., McKernan K., Talamas J., Tirrell A., Ye W.,
RA   Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,
RA   Hedderich R., Ingram-Smith C., Kuettner H.C., Krzycki J.A.,
RA   Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
RA   Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,
RA   Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
RA   Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
RA   Metcalf W.W., Birren B.;
RT   "The genome of Methanosarcina acetivorans reveals extensive metabolic
RT   and physiological diversity.";
RL   Genome Res. 12:532-542(2002).
DR   EMBL; AE011160; AAM07744.1; -.
DR   InterPro; IPR001441; UPP_synth.
DR   Pfam; PF01255; UPP_synthetase; 1.
DR   ProDom; PD003461; UPP_synth; 1.
DR   TIGRFAMs; TIGR00055; uppS; 1.
KW   Complete proteome.
SQ   SEQUENCE   224 AA;  25704 MW;  D0BEB1FC71DE9607 CRC64;

  Query Match           76.7%;  Score 33;  DB 17;  Length 224;
  Best Local Similarity  77.8%;  Pred. No. 40;
  Matches    7;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 FTLASSETG 9
              ||  ||||:
Db         92 FTCLSSETG 100


RESULT 8
Q8SPQ0
ID   Q8SPQ0      PRELIMINARY;      PRT;   383 AA.
AC   Q8SPQ0;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   BP40 precursor.
OS   Capra hircus (Goat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Caprinae; Capra.
OX   NCBI_TaxID=9925;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Mammary gland;
RA   Paramasivam M., Saravanan K., Sharma S., Mohanty A.K., Yadav S.,
RA   Srinivasan A., Singh T.P.;
RT   "Goat mammary gland mRNA for a novel protein, BP40, complete coding
RT   region.";
RL   Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY081150; AAL87007.1; -.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   ProDom; PD000471; Glyco_hydro_18; 1.
DR   SMART; SM00636; Glyco_18; 1.
KW   Signal.
FT   SIGNAL        1     16       POTENTIAL.
SQ   SEQUENCE   383 AA;  42893 MW;  17655ED4BE4E9F5F CRC64;

  Query Match           76.7%;  Score 33;  DB 6;  Length 383;
  Best Local Similarity  87.5%;  Pred. No. 73;
  Matches    7;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTLASSET 8
              |:|||||:
Db        265 FTLASSKT 272


RESULT 9
Q29411
ID   Q29411      PRELIMINARY;      PRT;   383 AA.
AC   Q29411;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   38 kDa heparin-binding glycoprotein.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Smooth muscle;
RX   MEDLINE=95286589; PubMed=7768902;
RA   Shackelton L.M., Mann D.M., Millis A.J.;
RT   "Identification of a 38-kDa heparin-binding glycoprotein (gp38k) in
RT   differentiating vascular smooth muscle cells as a member of a group of
RT   proteins associated with tissue remodeling.";
RL   J. Biol. Chem. 270:13076-13083(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Smooth muscle;
RA   Shackelton L.M., Mann D.M., Millis A.J.T.;
RL   Submitted (JAN-1995) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U19900; AAA86482.1; -.
DR   EMBL; Z47803; CAA87764.1; -.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   ProDom; PD000471; Glyco_hydro_18; 1.
DR   SMART; SM00636; Glyco_18; 1.
KW   Heparin-binding.
SQ   SEQUENCE   383 AA;  42443 MW;  3D1039F49910BDC4 CRC64;

  Query Match           76.7%;  Score 33;  DB 6;  Length 383;
  Best Local Similarity  87.5%;  Pred. No. 73;
  Matches    7;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 FTLASSRT 8
              ||||||:|
Db        265 FTLASSKT 272


RESULT 10
Q9P2G7
ID   Q9P2G7      PRELIMINARY;      PRT;  1265 AA.
AC   Q9P2G7;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical protein KIAA1380 (Fragment).
GN   KIAA1380.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=20181126; PubMed=10718198;
RA   Nagase T., Kikuno R., Ishikawa K., Hirosawa M., Ohara O.;
RT   "Prediction of the coding sequences of unidentified human genes.XVI.
RT   The complete sequences of 150 new cDNA clones from brain which code
RT   for large proteins in vitro.";
RL   DNA Res. 7:65-73(2000).
DR   EMBL; AB037801; BAA92618.1; -.
DR   InterPro; IPR003347; TF_JmjC.
DR   Pfam; PF02373; jmjC; 1.
DR   SMART; SM00558; JmjC; 1.
KW   Hypothetical protein.
FT   NON_TER       1       1
SQ   SEQUENCE   1265 AA;  141291 MW;  33DB5BE53A7D9EBE CRC64;

  Query Match            76.7%;  Score 33;  DB 4;  Length 1265;
  Best Local Similarity  87.5%;  Pred. No. 2.8e+02;
  Matches    7;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 TLASSETG 9
              |:|||:||
Db        183 TLASSKTG 190


RESULT 11
Q8N3U0
ID   Q8N3U0      PRELIMINARY;      PRT;  2162 AA.
AC   Q8N3U0;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical protein (Fragment).
GN   DKFZP761F0118.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Amygdala;
RA   Ottenwaelder B., Obermaier B., Mewes H.W., Weil B., Wiemann S.;
RL   Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL831917; CAD38578.1; -.
DR   InterPro; IPR003347; TF_JmjC.
DR   Pfam; PF02373; jmjC; 1.
DR   SMART; SM00558; JmjC; 1.
KW   Hypothetical protein.
FT   NON_TER       1       1
SQ   SEQUENCE   2162 AA;  240377 MW;  BFE9F3F68A957108 CRC64;

  Query Match            76.7%;  Score 33;  DB 4;  Length 2162;
  Best Local Similarity  87.5%;  Pred. No. 5.1e+02;
  Matches    7;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 TLASSETG 9
              |:|||:||
Db       1079 TLASSKTG 1086


RESULT 12
O27403
ID   O27403      PRELIMINARY;      PRT;  251 AA.
AC   O27403;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Conserved protein.
GN   MTH1349.
OS   Methanobacterium thermoautotrophicum.
OC   Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC   Methanobacteriaceae; Methanothermobacter.
OX   NCBI_TaxID=187420;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Delta H;
RX   MEDLINE=98037514; PubMed=9371463;
RA   Smith D.R., Doucette-Stamm L.A., DeLoughery C., Lee H.-M., Dubois J.,
RA   Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RA   Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,
RA   Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA   Jiwani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA   McDougall S., Shimer G., Goyal A., Pietrovski S., Church G.M.,
RA   Daniels C.J., Mao J.-I., Rice P., Nelling J., Reeve J.N.;
RT   "Complete genome sequence of Methanobacterium thermoautotrophicum
RT   deltaH: functional analysis and comparative genomics.";
RL   J. Bacteriol. 179:7135-7155(1997).
CC   -!- SIMILARITY: BELONGS TO THE ARSR FAMILY OF TRANSCRIPTIONAL
CC       REGULATORS.
DR   EMBL; AE000898; AAB85826.1; -.
DR   InterPro; IPR001845; HTH_ArsR.
DR   InterPro; IPR004096; V4R.
DR   Pfam; PF01022; HTH_5; 1.
DR   Pfam; PF02830; V4R; 1.
DR   SMART; SM00418; HTH_ARSR; 1.
KW   DNA-binding; Transcription regulation; Complete proteome.
SQ   SEQUENCE   251 AA;  28923 MW;  51668971AF84BB2D CRC64;

  Query Match            74.4%;  Score 32;  DB 17;  Length 251;
  Best Local Similarity  66.7%;  Pred. No. 76;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FTLASSETG 9
              ||: |:|||
Db          7 FTIFSTETG 15


RESULT 13
Q9HPA0
ID   Q9HPA0      PRELIMINARY;      PRT;  261 AA.
AC   Q9HPA0;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Vng1735c.
GN   VNG1735C.
OS   Halobacterium sp. (strain NRC-1).
OC   Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC   Halobacteriaceae; Halobacterium.
OX   NCBI_TaxID=64091;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20504483; PubMed=11016950;
RA   Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
RA   Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Sbrogna J.,
RA   Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
RA   Leithauser B., Keller K., Cruz R., Danson M.J., Hough D.W.,
```

```
RA   Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
RA   Isenbarger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,
RA   Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA   Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S.;
RT   "Genome sequence of Halobacterium species NRC-1.";
RL   Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
DR   EMBL; AE005079; AAG19970.1; -.
DR   InterPro; IPR005302; MOSC.
DR   InterPro; IPR005303; MOSC_N.
DR   Pfam; PF03473; MOSC; 1.
DR   Pfam; PF03476; MOSC_N; 1.
KW   Complete proteome.
SQ   SEQUENCE   261 AA;  28217 MW;  6F5304AB9F778789 CRC64;

  Query Match             74.4%;  Score 32;  DB 17;  Length 261;
  Best Local Similarity   66.7%;  Pred. No. 79;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 FTLASSETG 9
              |||  ::||
Db         37 FTLVDSDTG 45


RESULT 14
Q9WTV1
ID   Q9WTV1      PRELIMINARY;      PRT;   352 AA.
AC   Q9WTV1;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Glycoprotein-39 (Fragment).
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
CX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Lewis;
RA   Wendling U., Boots A.M.H., van Eden W.;
RT   "Cloning of the rat homologue of Human Cartilage glycoprotein-39 a
RT   potential autoantigen in arthritis.";
RL   Submitted (APR-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF062038; AAD22610.1; -.
DR   HSSP; P07254; 1CTN.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   ProDom; PD000471; Glyco_hydro_18; 1.
DR   SMART; SM00636; Glyco_18; 1.
FT   NON_TER       1      1
FT   NON_TER     352    352
SQ   SEQUENCE   352 AA;  39391 MW;  CBDE991610AC936C CRC64;

  Query Match             74.4%;  Score 32;  DB 11;  Length 352;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FTLASSE 7
              |||||||
Db        244 FTLASSE 250


RESULT 15
Q90Y50
ID   Q90Y50      PRELIMINARY;      PRT;   372 AA.
AC   Q90Y50;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Coxsackievirus and adenovirus receptor-like protein.
GN   CXADR OR CAR.
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20535037; PubMed=11080637;
RA   van Raaij M.J., Chouin E., van der Zandt H., Bergelson J.M.,
RA   Cusack S.;
RT   "Dimeric structure of the coxsackievirus and adenovirus receptor D1
RT   domain at 1.7 A resolution.";
RL   Structure 8:1147-1155(2000).
DR   EMBL; AF268197; AAK58592.1; -.
DR   ZFIN; ZDB-GENE-020814-2; cxadr.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003006; Ig_MHC.
DR   Pfam; PF00047; ig; 2.
DR   SMART; SM00409; IG; 2.
DR   PROSITE; PS50835; IG_LIKE; 2.
KW   Receptor.
SQ   SEQUENCE   372 AA;  40664 MW;  C363B71E7601C73A CRC64;

  Query Match             74.4%;  Score 32;  DB 19;  Length 372;
  Best Local Similarity   66.7%;  Pred. No. 1.2e+02;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FTLASSETG 9
              |||:| ::|
Db         47 FTLASDDSG 55


Search completed: September 24, 2003, 13:47:22
Job time : 30.2247 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 24, 2003, 13:25:57 ; Search time 36.1011 Seconds
(without alignments)
39.570 Million cell updates/sec

Title: US-09-744-282-3
Perfect score: 46
Sequence: 1 YDDQFSVKS 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	46	100.0	9	17	AAR95073	Human cartilage gl
2	46	100.0	9	21	AAY69228	Peptide fragment o
3	46	100.0	13	17	AAR95077	Peptide contg. hC
4	46	100.0	13	21	AAY69232	Peptide fragment o
5	46	100.0	13	23	AAE20225	Human cartilage gl
6	46	100.0	16	18	AAW45150	Human cartilage gl
7	46	100.0	16	18	AAW45151	Human cartilage gl
8	46	100.0	105	22	AAU02074	Synthetic human ta
9	46	100.0	220	22	AAU02078	Synthetic human mu
10	46	100.0	281	22	AAU02079	Synthetic multi-ta
11	46	100.0	321	9	AAP81342	Polypeptide involv
12	46	100.0	350	22	AAU02076	Synthetic human mu
13	46	100.0	383	23	ABG96297	Human ovarian canc
14	46	100.0	383	24	ABU56651	Lung cancer-associ
15	46	100.0	459	22	AAU02077	Synthetic multi-ta
16	43	93.5	383	18	AAW26751	Bovine whey protei
17	43	93.5	383	21	AAB03442	Gp38k protein sequ
18	42	91.3	14	22	ABB55896	Vascular dementia-
19	42	91.3	14	22	ABB55940	Vascular dementia-
20	42	91.3	14	22	ABB56115	Vascular dementia-
21	42	91.3	14	22	ABB52309	Human API-41 trypt
22	42	91.3	14	22	AAU28537	DPI tryptic digest
23	42	91.3	14	22	AAU28750	DPI tryptic digest
24	42	91.3	14	22	AAU26184	Depression-Associa
25	42	91.3	14	22	AAU26398	Depression-Associa
26	36	78.3	702	23	ABB77732	Amino acid sequenc
27	35	76.1	235	24	ABP81382	Streptococcus pneu
28	35	76.1	235	24	ABU61414	S. pneumoniae type
29	34	73.9	216	22	AAB63816	Human prostate can
30	34	73.9	1139	22	ABG22397	Novel human diagno
31	33	71.7	85	21	AAY56193	Protein encoded by
32	33	71.7	106	22	AAG82259	S. epidermidis ope
33	33	71.7	119	23	ABP40383	Staphylococcus epi
34	33	71.7	123	20	AAY37811	Chlamydia trachoma
35	33	71.7	1228	21	AAG30727	Arabidopsis thalia
36	33	71.7	1230	21	AAG30726	Arabidopsis thalia
37	33	71.7	1248	21	AAG30725	Arabidopsis thalia
38	32	69.6	244	22	AAG81746	S. epidermidis ope
39	32	69.6	246	23	ABP40447	Staphylococcus epi
40	32	69.6	452	15	AAR63186	Variant alpha amyl
41	32	69.6	468	13	AAR24136	Alpha-amylase vari
42	32	69.6	468	15	AAR63187	Variant alpha amyl
43	32	69.6	468	15	AAR63184	Variant alpha amyl
44	32	69.6	469	15	AAR63185	Variant alpha amyl
45	32	69.6	494	8	AAP70577	Alpha-amylase gene

ALIGNMENTS

RESULT 1
AAR95073
ID AAR95073 standard; peptide; 9 AA.
XX
AC AAR95073;
XX
DT 17-DEC-1996 (first entry)
XX
DE Human cartilage glycoprotein 39 autoantigen derived peptide (ITI).
XX
KW Human; cartilage; glycoprotein 39; hC gp-39; autoantigen;
KW induction; T-cell; tolerance; mediated; destruction;
KW rheumatoid arthritis; disease model; diagnosis.
XX
OS Homo sapiens.
XX
PN WO9613517-A1.
XX
PD 09-MAY-1996.
XX
PF 25-OCT-1995; 95WO-EP04201.
XX
PR 07-APR-1995; 95EP-0200886.
PR 27-OCT-1994; 94EP-0203128.
XX
PA (ALKU) AKZO NOBEL NV.
XX
PI Boots AMH, Verheijden GFM;
XX
DR WPI; 1996-239447/24.
XX
PT Peptide(s) derived from human cartilage gp-39 autoantigen - used

PT to induce specific T cell tolerance, esp. for treatment of
PT rheumatoid arthritis, also useful in producing arthritic animal
PT models
XX
PS Claim 1; Page 28; 37pp; English.
XX
CC The present peptide, derived from the human cartilage glycoprotein
CC 39 (hC gp-39) autoantigen, can be used to induce specific T-cell
CC tolerance to hC gp-39 in a patient with T-cell mediated cartilage
CC destruction, esp. rheumatoid arthritis (RA). It may also be used to
CC induce RA in mice, etc. (human disease models), and to diagnose the
CC presence of activated T-cells autoreactive with hC gp-39, i.e. to
CC identify patients suitable for tolerising treatment.
CC Therapeutically the peptide is given in a dose of 0.1-100 microg/kg
CC by injection, or 10-50 microg/kg to produce a disease model. The
CC peptide has a specific tolerising effect on autoreactive T-cells,
CC but leaves other components of the immune system intact, and
CC therefore should have few side effects. In an assay to determine
CC the peptide's binding to HLA-DR mols., specifically DR4Dw4 (known
CC to be associated with an increased risk of RA development), the
CC peptide gave a DOT blot IC(50) (micromolar) value of 0.23 compared
CC to 0.55 for the control peptide HA(307-319)F.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 46; DB 17; Length 9;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YDDQESVKS 9
||||||||| |
Db 1 YDDQESVKS 9

RESULT 2
AAY69228
ID AAY69228 standard; peptide; 9 AA.
XX
AC AAY69228;
XX
DT 30-MAY-2000 (first entry)
XX
DE Peptide fragment of human cartilage gp-39 (HC gp-39) protein.
XX
KW Human cartilage gp-39 protein; HC gp-39 protein; T-cell activator;
KW inflammatory disease; lymphocyte; antigen; bystander suppression;
KW antigen-specific T-cell tolerance; autoimmune disorder;
KW inflammatory disease; rheumatoid arthritis;
KW autoimmune inflammatory disease; autoantigen.
XX
OS Homo sapiens.
XX
PN WO200004917-A2.
XX
PD 03-FEB-2000.
XX
PF 19-JUL-1999; 99WO-EP05331.
XX
PR 23-JUL-1998; 98EP-0202471.
XX
PA (ALKU) AKZO NOBEL NV.
XX
PI Miltenburg AMM, Boots AMH;
XX
DR WPI; 2000-182537/16.
XX
PT Use of cartilage protein or peptides for treating inflammatory diseases
PT especially rheumatoid arthritis -
XX
PS Claim 2; Page 17; 29pp; English.
XX
CC AAY69228-33 represent fragments of the human cartilage gp-39 (HC gp-39)
CC protein. The peptides are T-cell activators. The peptides are used for
CC preventing inflammatory diseases, by modulating lymphocytes reactive
CC to antigens other than HC gp-39, present in the same tissue of
CC HC gp-39. By the induction of antigen-specific T-cell tolerance,
CC autoimmune disorders can be treated by bystander suppression. The
CC HC gp-39 peptides are useful for preventing inflammatory diseases
CC especially rheumatoid arthritis. The fragments may also be used for
CC treating autoimmune inflammatory diseases. The HC gp-39 peptides are
CC especially beneficial in autoimmune conditions in which the
CC autoantigens associated with the condition are unknown.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 46; DB 21; Length 9;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YDDQESVKS 9
||||||||| |
Db 1 YDDQESVKS 9

RESULT 3
AAR95077
ID AAR95077 standard; peptide; 13 AA.
XX
AC AAR95077;
XX
DT 17-DEC-1996 (first entry)
XX
DE Peptide contg. hC gp-39 autoantigen derived peptide (III).
XX
KW Human; cartilage; glycoprotein 39; HC gp-39; autoantigen;
KW induction; T-cell; tolerance; mediated; destruction;
KW rheumatoid arthritis; disease model; diagnosis.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 3..11
FT /note= "HC gp-39 peptide (III)"
XX
PN WO9613517-A1.
XX
PD 09-MAY-1996.
XX
PF 25-OCT-1995; 95WO-EP04201.
XX
PR 07-APR-1995; 95EP-0200886.
PR 27-OCT-1994; 94EP-0203128.
XX
PA (ALKO) AKZO NOBEL NV.
XX
PI Boots AMH, Verheijden GFM;
XX
DR WPI; 1996-239447/24.
XX
PT Peptide(s) derived from human cartilage gp-39 autoantigen - used
PT to induce specific T cell tolerance, esp. for treatment of
PT rheumatoid arthritis, also useful in producing arthritic animal
PT models
XX
PS Claim 2; Page 28; 37pp; English.
XX
CC The present peptide, contg. a peptide derived from the human
CC cartilage glycoprotein 39 (hC gp-39) autoantigen, can be used to
CC induce specific T-cell tolerance to hC gp-39 in a patient with
CC T-cell mediated cartilage destruction, esp. rheumatoid arthritis
CC (RA). It may also be used to induce RA in mice, etc. (human disease
CC models), and to diagnose the presence of activated T-cells
CC autoreactive with hC gp-39, i.e. to identify patients suitable for
CC tolerising treatment.
CC Therapeutically the peptide is given in a dose of 0.1-100 microg/kg
CC by injection, or 10-50 microg/kg to produce a disease model. The

```
CC   peptide has a specific tolerising effect on autoreactive T-cells,
CC   but leaves other components of the immune system intact, and
CC   therefore should have few side effects.
XX
SQ   Sequence   13 AA;

  Query Match             100.0%;  Score 46;  DB 17;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 0.017;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YDDQESVKS 9
              |||||||||
Db          3 YDDQESVKS 11


RESULT 4
AAY69232
ID   AAY69232 standard; peptide; 13 AA.
XX
AC   AAY69232;
XX
DT   30-MAY-2000  (first entry)
XX
DE   Peptide fragment of human cartilage gp-39 (HC gp-39) protein.
XX
KW   Human cartilage gp-39 protein; HC gp-39 protein; T-cell activator;
KW   inflammatory disease; lymphocyte; antigen; bystander suppression;
KW   antigen-specific T-cell tolerance; autoimmune disorder;
KW   inflammatory disease; rheumatoid arthritis;
KW   autoimmune inflammatory disease; autoantigen.
XX
OS   Homo sapiens.
XX
PN   WO200004917-A2.
XX
PD   03-FEB-2000.
XX
PF   19-JUL-1999;   99WO-EP05331.
XX
PR   23-JUL-1998;   98EP-0202471.
XX
PA   (ALKU ) AKZO NOBEL NV.
XX
PI   Miltenburg AMM,  Boots AMH;
XX
DR   WPI; 2000-182537/16.
XX
PT   Use of cartilage protein or peptides for treating inflammatory diseases
PT   especially rheumatoid arthritis  -
XX
PS   Claim 2; Page 17; 29pp; English.
XX
CC   AAY69226-33 represent fragments of the human cartilage gp-39 (HC gp-39)
CC   protein. The peptides are T-cell activators. The peptides are used for
CC   preventing inflammatory diseases, by modulating lymphocytes reactive
CC   to antigens other than HC gp-39, present in the same tissue of
CC   HC gp-39. By the induction of antigen-specific T-cell tolerance,
CC   autoimmune disorders can be treated by bystander suppression. The
CC   HC gp-39 peptides are useful for preventing inflammatory diseases
CC   especially rheumatoid arthritis. The fragments may also be used for
CC   treating autoimmune inflammatory diseases. The HC gp-39 peptides are
CC   especially beneficial in autoimmune conditions in which the
CC   autoantigens associated with the condition are unknown.
XX
SQ   Sequence   13 AA;

  Query Match             100.0%;  Score 46;  DB 21;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 0.017;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YDDQESVKS 9
              |||||||||
Db          3 YDDQESVKS 11


RESULT 5
AAE20225
ID   AAE20225 standard; peptide; 13 AA.
XX
AC   AAE20225;
XX
DT   18-JUN-2002  (first entry)
XX
DE   Human cartilage glycoprotein (HC gp)-39 peptide (residues 326-338).
XX
KW   Human cartilage glycoprotein-39; HC gp-39; autoimmune disorder; MHC;
KW   major histocompatibility complex; insulin-dependent diabetes mellitus;
KW   multiple sclerosis; psoriasis; myasthenia gravis; rheumatoid arthritis;
KW   immunosuppressive; neuroprotective; therapy.
XX
OS   Homo sapiens.
XX
PN   WO200214870-A2.
XX
PD   21-FEB-2002.
XX
PF   08-AUG-2001; 2001WO-EP09136.
XX
PR   14 AUG-2000; 2000EP-0202844.
XX
PA   (ALKU ) AKZO NOBEL NV.
XX
PI   Steenbakkers PGA;
XX
DR   WPI; 2002-269214/31.
XX
PT   Method for diagnosing an autoimmune disease activity by detecting the
PT   presence of an autoimmune specific major histocompatibility complex
PT   (MHC)-peptide complex in a patient using antibodies specific against
PT   MHC-peptide complexes -
XX
PS   Disclosure; Page 43; 48pp; English.
XX
CC   The patent discloses a method for diagnosing an autoimmune disease
CC   activity. The method involves detecting the presence of an autoimmune
CC   specific major histocompatibility complex (MHC)-peptide complex in a
CC   patient suffering from an autoimmune disease with antibodies or its
CC   antigen binding domains that specifically bind to a MHC-gp39-derived
CC   peptide. The method and antibodies are useful for diagnosing auto-
CC   immune diseases or purifying autoimmune specific MHC-peptide complexes.
CC   They are also useful in therapy or for the manufacture of pharmaceutical
CC   preparation for the treatment of autoimmune disorders such as insulin-
CC   dependent diabetes mellitus, multiple sclerosis, myasthenia gravis,
CC   psoriasis or rheumatoid arthritis. The present sequence is human
CC   cartilage glycoprotein (HC gp)-39 peptide (residues 326-338). This
CC   peptide is used in the exemplification of the invention.
XX
SQ   Sequence   13 AA;

  Query Match             100.0%;  Score 46;  DB 23;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 0.017;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YDDQESVKS 9
              |||||||||
Db          3 YDDQESVKS 11


RESULT 6
AAW45150
ID   AAW45150 standard; peptide; 16 AA.
XX
AC   AAW45150;
XX
DT   28-APR-1998  (first entry)
XX
```

```
DE   Human cartilage glycoprotein 39 derived peptide #51.
XX
KW   Articular cartilage; immunosuppressive therapy; antigen; autoantigen;
KW   immunological tolerance; T-cell; human cartilage glycoprotein 39;
KW   HC gp-39; rheumatoid arthritis; epitope.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9740068-A1.
XX
PD   30-OCT-1997.
XX
PF   22-APR-1997;   97WO-EP02051.
XX
PR   24-APR-1996;   96EP-0201106.
XX
PA   (ALKU ) AKZO NOBEL NV.
XX
PI   Boots AMH, Verheijden GFM;
XX
DR   WPI; 1997-535775/49.
XX
PT   Peptide suitable for use in antigen specific immunosuppressive
PT   therapy - resembles or mimics epitope present on HC gp-39, so
PT   inducing systemic immunological tolerance to rheumatoid arthritis
PT   auto:antigen
XX
PS   Disclosure; Page 17; 82pp; English.
XX
CC   The present sequence represents a peptide which resembles or mimics an
CC   epitope present on human cartilage glycoprotein 39 (HC gp-39), an
CC   autoantigen in rheumatoid arthritis. The invention relates to peptides
CC   consisting of 16-55 amino acid residues comprising at least one of the
CC   following 19 sequences: LVCYYTSWS; FLCTHIIYS; IIYSFANIS; LKTLLSVGG;
CC   FIKSVPPFL; FDGLDLAWL; LYPGRRDKQ; YDIAKISQH; LDFISIMTY; FISIMIYDF;
CC   FRGQEDASP; YAVGYMLRL; MLRLGAPAS; LAYYEICDF; LRGATVHRT; YLKDRQLAG;
CC   LAGAMVWAL; VWALDLDDF; or LDLDDFQGS. They can be used medically in
CC   antigen specific immunosuppressive therapy, particularly the treatment
CC   of T-cell mediated destruction of articular cartilage in autoimmune
CC   diseases (e.g. rheumatoid arthritis). They can also be used to detect
CC   activated autoreactive T cells in an individual. The peptides have a
CC   specific effect on the autoreactive T cells, thus leaving the other
CC   components of the immune system intact, unlike the non-specific
CC   suppressive effect of immunosuppressive drugs, and do not cause toxic
CC   side effects. The peptides are predominantly recognised by autoreactive
CC   T cells from rheumatoid arthritis patients, but rarely by those from
CC   healthy donors.
XX
SQ   Sequence   16 AA;

  Query Match             100.0%;  Score 46;  DB 18;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 0.021;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YDDQESVKS 9
              |||||||||
Db          7 YDDQESVKS 15


RESULT 7
AAW45151
ID   AAW45151 standard; peptide; 16 AA.
XX
AC   AAW45151;
XX
DT   28-APR-1998  (first entry)
XX
DE   Human cartilage glycoprotein 39 derived peptide #52.
XX
KW   Articular cartilage; immunosuppressive therapy; antigen; autoantigen;
KW   immunological tolerance; T-cell; human cartilage glycoprotein 39;
KW   HC gp-39; rheumatoid arthritis; epitope.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9740068-A1.
XX
PD   30-OCT-1997.
XX
PF   22-APR-1997;   97WO-EP02051.
XX
PR   24-APR-1996;   96EP-0201106.
XX
PA   (ALKU ) AKZO NOBEL NV.
XX
PI   Boots AMH, Verheijden GFM;
XX
DR   WPI; 1997-535775/49.
XX
PT   Peptide suitable for use in antigen specific immunosuppressive
PT   therapy - resembles or mimics epitope present on HC gp-39, so
PT   inducing systemic immunological tolerance to rheumatoid arthritis
PT   auto:antigen
XX
PS   Disclosure; Page 17; 82pp; English.
XX
CC   The present sequence represents a peptide which resembles or mimics an
CC   epitope present on human cartilage glycoprotein 39 (HC gp-39), an
CC   autoantigen in rheumatoid arthritis. The invention relates to peptides
CC   consisting of 16-55 amino acid residues comprising at least one of the
CC   following 19 sequences: LVCYYTSWS; FLCTHIIYS; IIYSFANIS; LKTLLSVGG;
CC   FIKSVPPFL; FDGLDLAWL; LYPGRRDKQ; YDIAKISQH; LDFISIMTY; FISIMIYDF;
CC   FRGQEDASP; YAVGYMLRL; MLRLGAPAS; LAYYEICDF; LRGATVHRT; YLKDRQLAG;
CC   LAGAMVWAL; VWALDLDDF; or LDLDDFQGS. They can be used medically in
CC   antigen specific immunosuppressive therapy, particularly the treatment
CC   of T-cell mediated destruction of articular cartilage in autoimmune
CC   diseases (e.g. rheumatoid arthritis). They can also be used to detect
CC   activated autoreactive T cells in an individual. The peptides have a
CC   specific effect on the autoreactive T cells, thus leaving the other
CC   components of the immune system intact, unlike the non-specific
CC   suppressive effect of immunosuppressive drugs, and do not cause toxic
CC   side effects. The peptides are predominantly recognised by autoreactive
CC   T cells from rheumatoid arthritis patients, but rarely by those from
CC   healthy donors.
XX
SQ   Sequence   16 AA;

  Query Match             100.0%;  Score 46;  DB 18;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 0.021;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YDDQESVKS 9
              |||||||||
Db          1 YDDQESVKS 9


RESULT 8
AAD02074
ID   AAD02074 standard; Protein; 105 AA.
XX
AC   AAD02074;
XX
DT   07-SEP-2001  (first entry)
XX
DE   Synthetic human target autoantigen shGLP/RA.
XX
KW   shGLP/RA; immunogenic epitope cluster; TEC;
KW   synthetic human chondrocyte glycoprotein 69/cartilage link protein;
KW   autoantigen; autoimmune disease; multiple sclerosis;
KW   rheumatoid arthritis; insulin-dependent diabetes mellitus;
KW   myasthenia gravis; uveitis; autoimmune hepatitis; thyroiditis; insulitis;
KW   orchitis; idiopathic thrombocytopenic purpura; inflammatory disease;
KW   Crohn's disease; ulcerative colitis.
XX
```

OS Synthetic.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 103
FT /label= OTHER
FT /note= "Other= STOP"
XX
PN WO200131037-A2.
XX
PD 03-MAY-2001.
XX
PF 26-OCT-2000; 2000WO-IL00688.
XX
PR 27-OCT-1999; 991L-0132611.
XX
PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Ben-Nun A, Kerlero De Rosbo N, Sappler GP;
XX
DR WPI; 2001-300515/31.
DR N-PSDB; AAS04827.
XX
PT Novel synthetic human target autoantigen gene useful for treating
PT autoimmune diseases such as multiple sclerosis, insulin-dependent
PT diabetes mellitus, rheumatoid arthritis, myasthenia gravis, and uveitis
PT -
XX
PS Claim 173; Fig 51; 182pp; English.
XX
CC The sequence represents shGLP/RA (synthetic human chondrocyte
CC glycoprotein 69/cartilage link protein) containing immunogenic epitope
CC clusters (IEC) from chondrocyte glycoprotein 69 and cartilage link
CC protein. The synthetic human target autoantigen genes of the invention
CC comprise sequences coding for at least 2 IECs of autoantigen(s) related
CC to a specific autoimmune disease. The synthetic human target autoantigen
CC genes are useful for treating autoimmune diseases such as multiple
CC sclerosis, insulin-dependent diabetes mellitus, rheumatoid arthritis,
CC myasthenia gravis, uveitis, autoimmune hepatitis, thyroiditis, insulitis,
CC orchitis, idiopathic thrombocytopenic purpura, and inflammatory diseases
CC (Crohn's disease, ulcerative colitis). The synthetic human target
CC autoantigen genes are also useful for diagnosis and/or monitoring the
CC progression of the autoimmune disease.
XX
SQ Sequence 105 AA;

Query Match 100.0%; Score 46; DB 22; Length 105;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YDDQESVKS 9
||||||||| |
Db 47 YDDQESVKS 55

RESULT 9
AAU02078
ID AAU02078 standard; Protein; 220 AA.
XX
AC AAU02078;
XX
DT 07-SEP-2001 (first entry)
XX
DE Synthetic human multi-target autoantigen Y-RAPc.
XX
KW Y-RAPc; immunogenic epitope cluster; IEC;
KW autoantigen; shCollagen/RA; shAggrecan/RA; shGLP/RA;
KW autoimmune disease; multiple sclerosis; rheumatoid arthritis;
KW insulin-dependent diabetes mellitus; myasthenia gravis; uveitis;
KW autoimmune hepatitis; thyroiditis; insulitis; orchitis;
KW idiopathic thrombocytopenic purpura; inflammatory disease;
KW Crohn's disease; ulcerative colitis; shMultiTAG.
XX

OS Synthetic.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 218
FT /label= OTHER
FT /note= "Other= STOP"
XX
PN WO200131037-A2.
XX
PD 03-MAY-2001.
XX
PF 26-OCT-2000; 2000WO-IL00688.
XX
PR 27-OCT-1999; 991L-0132611.
XX
PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Ben-Nun A, Kerlero De Rosbo N, Sappler GP;
XX
DR WPI; 2001-300515/31.
DR N-PSDB; AAS04839.
XX
PT Novel synthetic human target autoantigen gene useful for treating
PT autoimmune diseases such as multiple sclerosis, insulin-dependent
PT diabetes mellitus, rheumatoid arthritis, myasthenia gravis, and uveitis
PT -
XX
PS Claim 185; Fig 59; 182pp; English.
XX
CC The sequence represents synthetic human multi-target autoantigen
CC (shMultiTAG) Y-RAPa consisting of regions encoding preferred epitopes
CC from shCollagen/RA, shAggrecan/RA and shGLP/RA. The synthetic human
CC target autoantigen genes of the invention comprise sequences coding for
CC at least 2 IECs of autoantigen(s) related to a specific autoimmune
CC disease. The synthetic human target autoantigen genes are useful for
CC treating autoimmune diseases such as multiple sclerosis, insulin-
CC dependent diabetes mellitus, rheumatoid arthritis, myasthenia gravis,
CC uveitis, autoimmune hepatitis, thyroiditis, insulitis, orchitis,
CC idiopathic thrombocytopenic purpura, and inflammatory diseases (Crohn's
CC disease, ulcerative colitis). The synthetic human target autoantigen
CC genes are also useful for diagnosis and/or monitoring the progression of
CC the autoimmune disease.
XX
SQ Sequence 220 AA;

Query Match 100.0%; Score 46; DB 22; Length 220;
Best Local Similarity 100.0%; Pred. No. 0.42;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YDDQESVKS 9
||||||||| |
Db 203 YDDQESVKS 211

RESULT 10
AAU02079
ID AAU02079 standard; Protein; 281 AA.
XX
AC AAU02079;
XX
DT 07-SEP-2001 (first entry)
XX
DE Synthetic multi-target autoantigen Y-RAPd.
XX
KW Y-RAPd; immunogenic epitope cluster; IEC;
KW autoantigen; shCollagen/RA; shAggrecan/RA; shGLP/RA; shHSP/RA;
KW autoimmune disease; multiple sclerosis; rheumatoid arthritis;
KW insulin-dependent diabetes mellitus; myasthenia gravis; uveitis;
KW autoimmune hepatitis; thyroiditis; insulitis; orchitis;
KW idiopathic thrombocytopenic purpura; inflammatory disease;
KW Crohn's disease; ulcerative colitis; shMultiTAG.
XX

```
OS   Synthetic.
OS   Homo sapiens.
OS   Escherichia coli.
OS   Mycobacterium tuberculosis.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 279
FT                   /label= OTHER
FT                   /note= "Other= STOP"
XX
PN   WO200131037-A2.
XX
PD   03-MAY-2001.
XX
PF   26-OCT-2000; 2000WO-IL00688.
XX
PR   27-OCT-1999;   99IL-0132611.
XX
PA   (YEDA ) YEDA RES & DEV CO LTD.
XX
PI   Ben-Nun A,  Kerlero De Rosbo N,  Sappler GP;
XX
DR   WPI; 2001-300515/31.
DR   N-PSDB; AAS04840.
XX
PT   Novel synthetic human target autoantigen gene useful for treating
PT   autoimmune diseases such as multiple sclerosis, insulin-dependent
PT   diabetes mellitus, rheumatoid arthritis, myasthenia gravis, and uveitis
PT   -
XX
PS   Claim 188; Fig 61; 182pp; English.
XX
CC   The sequence represents synthetic human multi-target autoantigen
CC   (shMultiTAG) Y-RAPd consisting of preferred epitopes encoded by
CC   shCollagen/RA, shAggrecan/RA, shGLP/RA and shHSP/RA. The synthetic human
CC   target autoantigen genes of the invention comprise sequences coding for
CC   at least 2 IECs of autoantigen(s) related to a specific autoimmune
CC   disease. The synthetic human target autoantigen genes are useful for
CC   treating autoimmune diseases such as multiple sclerosis, insulin-
CC   dependent diabetes mellitus, rheumatoid arthritis, myasthenia gravis,
CC   uveitis, autoimmune hepatitis, thyroiditis, insulitis, orchitis,
CC   idiopathic thrombocytopenic purpura, and inflammatory diseases (Crohn's
CC   disease, ulcerative colitis). The synthetic human target autoantigen
CC   genes are also useful for diagnosis and/or monitoring the progression of
CC   the autoimmune disease.
XX
SQ   Sequence   281 AA;

  Query Match             100.0%;  Score 46;  DB 22;  Length 281;
  Best Local Similarity   100.0%;  Pred. No. 0.55;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YDDQESVKS 9
              |||||||||
Db        203 YDDQESVKS 211


RESULT 11
AAP81342
ID   AAP81342 standard; protein; 321 AA.
XX
AC   AAP81342;
XX
DT   25-MAR-2003  (updated)
DT   10-MAR-2003  (updated)
DT   19-OCT-1990  (first entry)
XX
DE   Polypeptide involved in protective mechanisms.
XX
KW   Immune response; cell growth.
XX
OS   Unidentified.
XX
PN   JP63023898-A.
XX
PD   30-JAN-1988.
XX
PF   16-JUL-1986;   86JP-0167518.
XX
PR   16-JUL-1986;   86JP-0167518.
XX
PA   (DAIN ) DAINIPPON PHARM KK.
XX
DR   WPI; 1988-068419/10.
DR   P-PSDB; AAP81342.
XX
PT   New polypeptide and DNA encoding it - related to protective
PT   mechanisms such as immune response etc.
XX
PS   Disclosure; Page ?; ?pp; Japanese.
XX
CC   This polypeptide is involved in protective mechanisms such as immune
CC   response, cell growth and activation of protective functions.
CC   (Updated on 10-MAR-2003 to add missing OS field.)
CC   (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence   321 AA;

  Query Match             100.0%;  Score 46;  DB 9;  Length 321;
  Best Local Similarity   100.0%;  Pred. No. 0.64;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YDDQESVKS 9
              |||||||||
Db        266 YDDQESVKS 274


RESULT 12
AAU02076
ID   AAU02076 standard; Protein; 350 AA.
XX
AC   AAU02076;
XX
DT   07-SEP-2001  (first entry)
XX
DE   Synthetic human multi-target autoantigen Y-RAPa.
XX
KW   Y-RAPa; immunogenic epitope cluster; IEC;
KW   autoantigen; shCollagen/RA; shAggrecan/RA; shGLP/RA;
KW   autoimmune disease; multiple sclerosis; rheumatoid arthritis;
KW   insulin-dependent diabetes mellitus; myasthenia gravis; uveitis;
KW   autoimmune hepatitis; thyroiditis; insulitis; orchitis;
KW   idiopathic thrombocytopenic purpura; inflammatory disease;
KW   Crohn's disease; ulcerative colitis; shMultiTAG.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 348
FT                   /label= OTHER
FT                   /note= "Other= STOP"
XX
PN   WO200131037-A2.
XX
PD   03-MAY-2001.
XX
PF   26-OCT-2000; 2000WO-IL00688.
XX
PR   27-OCT-1999;   99IL-0132611.
XX
PA   (YEDA ) YEDA RES & DEV CO LTD.
XX
PI   Ben-Nun A,  Kerlero De Rosbo N,  Sappler GP;
XX
DR   WPI; 2001-300515/31.
```

DR N-PSDB; AAS04837.
XX
PT Novel synthetic human target autoantigen gene useful for treating
PT autoimmune diseases such as multiple sclerosis, insulin-dependent
PT diabetes mellitus, rheumatoid arthritis, myasthenia gravis, and uveitis
PT -
XX
PS Claim 179; Fig 55; 182pp; English.
XX
CC The sequence represents synthetic human multi-target autoantigen
CC (shMultiTAG) Y-RAPa consisting of regions encoding shCollagen/RA,
CC shAggrecan/RA and shGLP/RA. The synthetic human target autoantigen genes
CC of the invention comprise sequences coding for at least 2 IECs of
CC autoantigen(s) related to a specific autoimmune disease. The synthetic
CC human target autoantigen genes are useful for treating autoimmune
CC diseases such as multiple sclerosis, insulin-dependent diabetes
CC mellitus, rheumatoid arthritis, myasthenia gravis, uveitis, autoimmune
CC hepatitis, thyroiditis, insulitis, orchitis, idiopathic thrombocytopenic
CC purpura, and inflammatory diseases (Crohn's disease, ulcerative colitis).
CC The synthetic human target autoantigen genes are also useful for
CC diagnosis and/or monitoring the progression of the autoimmune disease.
XX
SQ Sequence 350 AA;

Query Match 100.0%; Score 46; DB 22; Length 350;
Best Local Similarity 100.0%; Pred. No. 0.71;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YDDQESVKS 9
||||||||| |
Db 292 YDDQESVKS 300

RESULT 13
ABG96297
ID ABG96297 standard; Protein; 383 AA.
XX
AC ABG96297;
XX
DT 11-DEC-2002 (first entry)
XX
DE Human ovarian cancer marker OV6.
XX
KW Human; ovarian cancer; marker; cancer; familial history; brain disorder;
KW central nervous system disorder; bacterial meningitis; viral meningitis;
KW Alzheimer's disease; Parkinson's disease; cerebral oedema; hydrocephalus;
KW brain herniation; inflammation; encephalitis; testicular disorder;
KW nontuberculous granulomatous orchitis; connective tissue disorder;
KW heart disorder; ischaemic heart disease; atherosclerosis; neoplasm;
KW histological type; carcinogenic; ovarian cancer marker.
XX
OS Homo sapiens.
XX
PN WO200271926-A2.
XX
PD 19-SEP-2002.
XX
PF 14-MAR-2002; 2002WO-US07826.
XX
PR 14-MAR-2001; 2001US-276025P.
PR 14-MAR-2001; 2001US-276026P.
PR 10-AUG-2001; 2001US-311732P.
PR 19-SEP-2001; 2001US-323580P.
PR 26-SEP-2001; 2001US-324967P.
PR 26-SEP-2001; 2001US-325102P.
PR 26-SEP-2001; 2001US-325149P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Monahan JE, Gannavarapu M, Hoersch S, Kamatkar S, Kovatis SG;
PI Meyers RE, Morrisey MP, Olandt PJ, Sen A, Vieby PO, Mills GB;
PI Bast RC, Lu K, Schmandt RE, Zhao X, Glatt K;
XX
DR WPI; 2002-723277/78.
DR N-PSDB; ABS76589.
XX
PT Assessing whether a patient is afflicted with ovarian cancer, useful in
PT assessing the stage or progression of the disease, comprises comparing
PT the expression level of a cancer marker in a sample from a patient and
PT from a non cancer patient -
XX
PS Disclosure; Page 164-165; 481pp; English.
XX
CC The present invention relates to a new method for assessing whether a
CC patient is afflicted with ovarian cancer. The method involves comparing
CC the expression level of a marker in a patient sample and the normal level
CC of expression of the marker in a control non-ovarian cancer sample, where
CC the marker is selected from 363 cancer markers described in the
CC specification. The method of the invention is useful in diagnosing or
CC characterising cancer, in detecting the presence of cancer as early as
CC possible, and the recurrence of ovarian cancer. The method may also be of
CC particular use with patients having an enhanced risk of developing
CC ovarian cancer (e.g. patients having a familial history of ovarian
CC cancer). The cancer markers may be used in the management and treatment
CC of e.g. brain and central nervous system disorders (e.g. bacterial and
CC viral meningitis, Alzheimer's disease or Parkinson's disease), brain
CC disorders (e.g. cerebral oedema, hydrocephalus or brain herniations),
CC inflammations (e.g. bacterial or viral meningitis or encephalitis),
CC testicular disorders (e.g. nontuberculous granulomatous orchitis),
CC connective tissue disorders, or heart disorders (e.g. ischaemic heart
CC disease or atherosclerosis). The compositions and methods may also be
CC used in assessing the histological type of neoplasm associated with
CC ovarian cancer, monitoring the progression of ovarian cancer,
CC determining whether ovarian cancer has metastasized or is likely to
CC metastasize, selecting a composition for inhibiting ovarian cancer,
CC assessing the ovarian carcinogenic potential of a compound, or
CC inhibiting ovarian cancer or at risk of developing ovarian cancer. The
CC present amino acid sequence represents one of the ovarian cancer markers
CC described in the invention.
XX
SQ Sequence 383 AA;

Query Match 100.0%; Score 46; DB 23; Length 383;
Best Local Similarity 100.0%; Pred. No. 0.78;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YDDQESVKS 9
||||||||| |
Db 328 YDDQESVKS 336

RESULT 14
AB056651
ID AB056651 standard; Protein; 383 AA.
XX
AC AB056651;
XX
DT 02-APR-2003 (first entry)
XX
DE Lung cancer-associated polypeptide #244.
XX
KW Lung cancer-associated polypeptide; cytostatic; emphysema;
KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
XX
OS Unidentified.
XX
PN WO200286443-A2.
XX
PD 31-OCT-2002.
XX
PF 18-APR-2002; 2002WO-US12476.
XX
PR 18-APR-2001; 2001US-284770P.

PR 10-MAY-2001; 2001US-290492P.
PR 09-NOV-2001; 2001US-339245P.
PR 13-NOV-2001; 2001US-350566P.
PR 29-NOV-2001; 2001US-334370P.
PR 12-APR-2002; 2002US-372246P.
XX
PA (EOSB-) ECS BIOTECHNOLOGY INC.
XX
PI Aziz N, Murray R;
XX
DR WPI; 2003-093161/08.
DR N-PSDB; ABX76380.
XX
PT Detecting a lung cancer-associated transcript in a cell from a patient
PT for treating lung cancer, by contacting a biological sample from the
PT patient with a polynucleotide that exhibits increased or decreased
PT expression in lung cancer -
XX
PS Claim 27; Page 377; 453pp; English.
XX
CC The invention relates to a method for detecting a lung cancer-associated
CC transcript in a cell from a patient, comprising contacting a biological
CC sample from the patient with a polynucleotide that selectively hybridises
CC to a sequence that is at least 80 % identical to a gene that exhibits
CC increased or decreased expression in lung cancer samples. Lung
CC cancer-associated polynucleotides and polypeptides are used for
CC identifying a compound that modulates a lung cancer-associated
CC polypeptide, for inhibiting proliferation of a lung cancer-associated
CC cell to treat lung cancer in a patient and for treating a mammal having
CC lung cancer by administering a modulatory compound identified. The
CC methods are useful for treating lung cancer, such as small cell lung
CC cancer, non-small cell lung cancer or other benign or precancerous
CC lesions, e.g. atelectasis, emphysema, bronchitis, chronic obstructive
CC pulmonary disease, fibrosis, hypersensitivity pneumonitis, interstitial
CC pulmonary fibrosis, asthma and bronchiectasis. The genes, polynucleotides
CC and polypeptides are useful for diagnostic purposes and as targets for
CC screening for therapeutic compounds that modulate lung cancer, such as
CC antibodies. Sequences ABU56408-ABU56745 represent lung cancer-associated
CC polypeptides of the invention.
XX
SQ Sequence 383 AA;

Query Match 100.0%; Score 46; DB 24; Length 383;
Best Local Similarity 100.0%; Pred. No. 0.78;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YDDQESVKS 9
||||||||| |
Db 328 YDDQESVKS 336

RESULT 15
AAU02077
ID AAU02077 standard; Protein; 459 AA.
XX
AC AAU02077;
XX
DT 07-SEP-2001 (first entry)
XX
DE Synthetic multi-target autoantigen Y-RAPb.
XX
KW Y-RAPa; immunogenic epitope cluster; IEC;
KW autoantigen; shCollagen/RA; shAggrecan/RA; shGLP/RA; shHSP/RA;
KW autoimmune disease; multiple sclerosis; rheumatoid arthritis;
KW insulin-dependent diabetes mellitus; myasthenia gravis; uveitis;
KW autoimmune hepatitis; thyroiditis; insulitis; orchitis;
KW idiopathic thrombocytopenic purpura; inflammatory disease;
KW Crohn's disease; ulcerative colitis; shMultiTAG.
XX
OS Synthetic.
OS Homo sapiens.
OS Escherichia coli.
OS Mycobacterium tuberculosis.

XX
FH Key Location/Qualifiers
FT Misc-difference 455
FT /label= OTHER
FT /note= "Other= STOP"
FT Misc-difference 457
FT /label= OTHER
FT /note= "Other= STOP"
XX
PN WO200131037-A2.
XX
PD 03-MAY-2001.
XX
PF 26-OCT-2000; 2000WO-IL00688.
XX
PR 27-OCT-1999; 99IL-0132611.
XX
PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Ben-Nun A, Kerlero De Rosbo N, Sappler GP;
XX
DR WPI; 2001-300515/31.
DR N-PSDB; AAS04838.
XX
PT Novel synthetic human target autoantigen gene useful for treating
PT autoimmune diseases such as multiple sclerosis, insulin-dependent
PT diabetes mellitus, rheumatoid arthritis, myasthenia gravis, and uveitis
PT -
XX
PS Claim 182; Fig 57; 182pp; English.
XX
CC The sequence represents synthetic human multi-target autoantigen
CC (shMultiTAG) Y-RAPa consisting of regions encoding shCollagen/RA,
CC shAggrecan/RA, shGLP/RA and shHSP/RA. The synthetic human target
CC autoantigen genes of the invention comprise sequences coding for at least
CC 2 IECs of autoantigen(s) related to a specific autoimmune disease. The
CC synthetic human target autoantigen genes are useful for treating
CC autoimmune diseases such as multiple sclerosis, insulin-dependent
CC diabetes mellitus, rheumatoid arthritis, myasthenia gravis, uveitis,
CC autoimmune hepatitis, thyroiditis, insulitis, orchitis, idiopathic
CC thrombocytopenic purpura, and inflammatory diseases (Crohn's disease,
CC ulcerative colitis). The synthetic human target autoantigen genes are
CC also useful for diagnosis and/or monitoring the progression of the
CC autoimmune disease.
XX
SQ Sequence 459 AA;

Query Match 100.0%; Score 46; DB 22; Length 459;
Best Local Similarity 100.0%; Pred. No. 0.96;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YDDQESVKS 9
||||||||| |
Db 292 YDDQESVKS 300

Search completed: September 24, 2003, 13:41:21
Job time : 37.2261 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 24, 2003, 13:35:23 ; Search time 10.9213 Seconds
(without alignments)
34.867 Million cell updates/sec

Title: US-09-744-282-3
Perfect score: 46
Sequence: 1 YDDQESVKS 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	46	100.0	9	1	US-08-619-645-3	Sequence 3, Appli
2	46	100.0	9	2	US-08-634-493-3	Sequence 3, Appli
3	46	100.0	13	1	US-08-619-645-7	Sequence 7, Appli
4	46	100.0	13	2	US-08-634-493-7	Sequence 7, Appli
5	46	100.0	16	3	US-09-171-705-51	Sequence 51, Appl
6	46	100.0	16	3	US-09-171-705-52	Sequence 52, Appl
7	43	93.5	383	4	US-09-459-749D-17	Sequence 17, Appl
8	33	71.7	119	4	US-09-134-001C-5228	Sequence 5228, Ap
9	32	69.6	246	4	US-09-134-001C-5292	Sequence 5292, Ap
10	32	69.6	468	1	US-08-204-656B-2	Sequence 2, Appli
11	32	69.6	468	1	US-08-204-656B-4	Sequence 4, Appli
12	32	69.6	468	1	US-08-204-656B-6	Sequence 6, Appli
13	32	69.6	468	1	US-08-204-656B-8	Sequence 8, Appli
14	32	69.6	468	1	US-08-470-702-6	Sequence 6, Appli
15	32	69.6	468	1	US-08-470-702-7	Sequence 7, Appli
16	32	69.6	468	1	US-08-470-702-8	Sequence 8, Appli
17	32	69.6	468	1	US-08-470-702-9	Sequence 9, Appli
18	32	69.6	468	1	US-08-467-831-6	Sequence 6, Appli
19	32	69.6	468	1	US-08-467-831-7	Sequence 7, Appli
20	32	69.6	468	1	US-08-467-831-8	Sequence 8, Appli
21	32	69.6	468	1	US-08-467-831-9	Sequence 9, Appli
22	31	67.4	188	2	US-08-933-750C-39	Sequence 39, Appl
23	31	67.4	188	3	US-09-234-613-39	Sequence 39, Appl
24	31	67.4	371	2	US-08-591-629-2	Sequence 2, Appli
25	31	67.4	1233	4	US-09-252-991A-23237	Sequence 23237, A
26	30	65.2	181	4	US-09-134-001C-3897	Sequence 3897, Ap
27	30	65.2	199	4	US-09-328-352-7220	Sequence 7220, Ap
28	30	65.2	244	4	US-09-328-352-4316	Sequence 4316, Ap
29	30	65.2	299	4	US-09-328-352-4624	Sequence 4624, Ap
30	30	65.2	349	4	US-09-328-352-7606	Sequence 7606, Ap
31	30	65.2	367	3	US-08-467-023-95	Sequence 95, Appl
32	30	65.2	449	1	US-08-074-121-6	Sequence 6, Appli
33	30	65.2	449	5	PCT-US94-06447-6	Sequence 6, Appli
34	30	65.2	465	4	US-09-252-991A-26980	Sequence 26980, A
35	30	65.2	489	1	US-07-903-103-4	Sequence 4, Appli
36	30	65.2	489	1	US-08-044-619A-4	Sequence 4, Appli
37	30	65.2	489	1	US-08-283-911-4	Sequence 4, Appli
38	30	65.2	489	1	US-08-245-500A-5	Sequence 5, Appli
39	30	65.2	489	1	US-08-390-546-5	Sequence 5, Appli
40	30	65.2	489	1	US-08-390-479A-5	Sequence 5, Appli
41	30	65.2	489	1	US-08-557-393-5	Sequence 5, Appli
42	30	65.2	489	1	US-08-390-516C-5	Sequence 5, Appli
43	30	65.2	489	1	US-08-390-517A-5	Sequence 5, Appli
44	30	65.2	489	1	US-08-390-515A-5	Sequence 5, Appli
45	30	65.2	489	2	US-08-801-728-5	Sequence 5, Appli

ALIGNMENTS

```
RESULT 1
US-08-619-645-3
; Sequence 3, Application US/08619645
; Patent No. 5736507
; GENERAL INFORMATION:
;   APPLICANT: Boots, Anna M.H.
;   APPLICANT: Verheijden, Gijsbertus F.M.
;   TITLE OF INVENTION: No. 5736507el peptides derived from
;   TITLE OF INVENTION: autoantigen for use in immunotherapy of autoimmune
;   TITLE OF INVENTION: disease
;   NUMBER OF SEQUENCES: 9
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Akzo No. 5736507el Patent Department
;     STREET: 1300 Piccard Drive, Suite 206
;     CITY: Rockville
;     STATE: Maryland
;     COUNTRY: U.S.A.
;     ZIP: 20850
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.25(EPO)
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/619,645
;     FILING DATE: 25-MAR-1996
;     CLASSIFICATION: 514
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Gormley, Mary E.
;     REGISTRATION NUMBER: 34,409
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (301) 947-4433
; INFORMATION FOR SEQ ID NO: 3:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 9 amino acids
;     TYPE: amino acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;   MOLECULE TYPE: peptide
US-08-619-645-3

  Query Match          100.0%;  Score 46;  DB 1;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 2.5e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YDDQESVKS 9
              |||||||||
Db          1 YDDQESVKS 9
```

```
RESULT 2
US-08-634-493-3
; Sequence 3, Application US/08634493
; Patent No. 5843449
;  GENERAL INFORMATION:
;    APPLICANT:  A.M.H. Boots
;    APPLICANT:  G.F.M. Verheijden
;    APPLICANT:  E.S. Bos
;    TITLE OF INVENTION:  No. 5843449el Peptides derived from autoantigen for use
;    TITLE OF INVENTION:  in Immunotherapy of Autoimmune Diseases
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Akzo No. 5843449el Patent Department
;      STREET:  1300 Piccard Drive, Suite 206
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/634,493
;      FILING DATE:  18-APR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/619,645
;      FILING DATE:  25-MAR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/EP95/04201
;      FILING DATE:  25-OCT-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  NL 942031287
;      FILING DATE:  27-OCT-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  NL 952008860
;      FILING DATE:  07-APR-1995
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 258-5200
;      TELEFAX:  (301) 977-0847
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-634-493-3

  Query Match           100.0%;  Score 46;  DB 2;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 2.5e-05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YDDQESVKS 9
              |||||||||
Db          1 YDDQESVKS 9


RESULT 3
US-08-619-645-7
; Sequence 7, Application US/08619645
; Patent No. 5736507
;  GENERAL INFORMATION:
;    APPLICANT:  Boots, Anna M.H.
;    APPLICANT:  Verheijden, Gijsbertus F.M.
;    TITLE OF INVENTION:  No. 5736507el peptides derived from
;    TITLE OF INVENTION:  autoantigen for use in immunotherapy of autoimmune
;    TITLE OF INVENTION:  disease
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Akzo No. 5736507el Patent Department
;      STREET:  1300 Piccard Drive, Suite 206
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  U.S.A.
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25(EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/619,645
;      FILING DATE:  25-MAR-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Gormley, Mary E.
;      REGISTRATION NUMBER:  34,409
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 947-4433
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  13 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-619-645-7

  Query Match           100.0%;  Score 46;  DB 1;  Length 13;
  Best Local Similarity  100.0%;  Pred. No. 0.0033;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YDDQESVKS 9
              |||||||||
Db          3 YDDQESVKS 11


RESULT 4
US-08-634-493-7
; Sequence 7, Application US/08634493
; Patent No. 5843449
;  GENERAL INFORMATION:
;    APPLICANT:  A.M.H. Boots
;    APPLICANT:  G.F.M. Verheijden
;    APPLICANT:  E.S. Bos
;    TITLE OF INVENTION:  No. 5843449el Peptides derived from autoantigen for use
;    TITLE OF INVENTION:  in Immunotherapy of Autoimmune Diseases
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Akzo No. 5843449el Patent Department
;      STREET:  1300 Piccard Drive, Suite 206
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/634,493
;      FILING DATE:  18-APR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/619,645
;      FILING DATE:  25-MAR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/EP95/04201
;      FILING DATE:  25-OCT-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  NL 942031287
;      FILING DATE:  27-OCT-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  NL 952008860
```

```
;      FILING DATE:  07-APR-1995
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 258-5200
;      TELEFAX:  (301) 977-0847
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  13 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-634-493-7

  Query Match             100.0%;  Score 46;  DB 2;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 0.0033;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YDDQESVKS 9
              |||||||||
Db          3 YDDQESVKS 11


RESULT 5
US-09-171-705-51
; Sequence 51, Application US/09171705
; Patent No. 6184204
; GENERAL INFORMATION:
;  APPLICANT: BOOTS, ANNA M.H.
;  APPLICANT: VERHEIJDEN, GILBERTUS F.M.
;  TITLE OF INVENTION: NOVEL PEPTIDES SUITABLE FOR USE IN ANTIGEN SPECIFIC
;  TITLE OF INVENTION: IMMUNOSUPPRESSIVE THERAPY
;  FILE REFERENCE: O/96198 US
;  CURRENT APPLICATION NUMBER: US/09/171,705
;  CURRENT FILING DATE: 1999-02-09
;  NUMBER OF SEQ ID NOS: 78
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 51
;   LENGTH: 16
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: DERIVED FROM
;   OTHER INFORMATION: SEQUENCE OF HUMAN CARTILAGE (HC) -39 PROTEIN
US-09-171-705-51

  Query Match             100.0%;  Score 46;  DB 3;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 0.0042;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YDDQESVKS 9
              |||||||||
Db          7 YDDQESVKS 15


RESULT 6
US-09-171-705-52
; Sequence 52, Application US/09171705
; Patent No. 6184204
; GENERAL INFORMATION:
;  APPLICANT: BOOTS, ANNA M.H.
;  APPLICANT: VERHEIJDEN, GILBERTUS F.M.
;  TITLE OF INVENTION: NOVEL PEPTIDES SUITABLE FOR USE IN ANTIGEN SPECIFIC
;  TITLE OF INVENTION: IMMUNOSUPPRESSIVE THERAPY
;  FILE REFERENCE: O/96198 US
;  CURRENT APPLICATION NUMBER: US/09/171,705
;  CURRENT FILING DATE: 1999-02-09
;  NUMBER OF SEQ ID NOS: 78
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 52
;   LENGTH: 16
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: DERIVED FROM
;   OTHER INFORMATION: SEQUENCE OF HUMAN CARTILAGE (HC) -39 PROTEIN
US-09-171-705-52

  Query Match             100.0%;  Score 46;  DB 3;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 0.0042;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YDDQESVKS 9
              |||||||||
Db          1 YDDQESVKS 9


RESULT 7
US-09-459-749D-17
; Sequence 17, Application US/09459749D
; Patent No. 6464975
; GENERAL INFORMATION:
;  APPLICANT: Millis, Albert J. T.
;  TITLE OF INVENTION: Compositions and Methods For Altering Cell Migration
;  FILE REFERENCE: 0794.016A
;  CURRENT APPLICATION NUMBER: US/09/459,749D
;  CURRENT FILING DATE:  1999-12-10
;  PRIOR APPLICATION NUMBER: 60/111,856
;  PRIOR FILING DATE: 1998-12-11
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
;   LENGTH: 383
;   TYPE: PRT
;   ORGANISM: Sus scrofa
US-09-459-749D-17

  Query Match             93.5%;  Score 43;  DB 4;  Length 383;
  Best Local Similarity   88.9%;  Pred. No. 0.53;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YDDQESVKS 9
              |||||||:|
Db        328 YDDQESVKN 336


RESULT 8
US-09-134-001C-5228
; Sequence 5228, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
;  APPLICANT: Lynn Doucette-Stamm et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCO
;  TITLE OF INVENTION:  EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: GTC-007
;  CURRENT APPLICATION NUMBER: US/09/134,001C
;  CURRENT FILING DATE:  1998-08-13
;  PRIOR APPLICATION NUMBER: US 60/064,964
;  PRIOR FILING DATE: 1997-11-08
;  PRIOR APPLICATION NUMBER: US 60/055,779
;  PRIOR FILING DATE: 1997-08-14
;  NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 5228
;   LENGTH: 119
;   TYPE: PRT
;   ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5228

  Query Match             71.7%;  Score 33;  DB 4;  Length 119;
  Best Local Similarity   62.5%;  Pred. No. 15;
  Matches    5;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YDDQESVK 8
              |||:|::
Db        109 YDDEESLR 116
```

```
RESULT 9
US-09-134-001C-5292
; Sequence 5292, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
;  APPLICANT: Lynn Doucette-Stamm et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
;  TITLE OF INVENTION:  EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: GTC-007
;  CURRENT APPLICATION NUMBER: US/09/134,001C
;  CURRENT FILING DATE:  1998-08-12
;  PRIOR APPLICATION NUMBER: US 60/064,964
;  PRIOR FILING DATE: 1997-11-08
;  PRIOR APPLICATION NUMBER: US 60/055,779
;  PRIOR FILING DATE: 1997-08-14
;  NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 5292
;   LENGTH: 246
;   TYPE: PRT
;   ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5292

  Query Match             69.6%;  Score 32;  DB 4;  Length 246;
  Best Local Similarity   66.7%;  Pred. No. 54;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 YDDQESVKS 9
              |||:  |||
Db         49 YDEQPDVKS 57


RESULT 10
US-08-204-656B-2
; Sequence 2, Application US/08204656B
; Patent No. 5538882
;  GENERAL INFORMATION:
;    APPLICANT:  Matsui, Ikuo
;    APPLICANT:  Ishikawa, Kazuhiko
;    APPLICANT:  Miyairi, Sachio
;    APPLICANT:  Honda, Koichi
;    TITLE OF INVENTION:  Variant-Type Carbohydrate Hydrolase,
;    TITLE OF INVENTION:  Variant Gene Of The Enzyme And Method For Producing
;    TITLE OF INVENTION:  Oligosaccharide Using The Enzyme
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Birch, Stewart, Kolasch & Birch, LLP
;      STREET:  8110 Gatehouse Road, Suite 500 East
;      CITY:  Falls Church
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22042
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/204,656B
;      FILING DATE:  02-MAR-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Weiner, Marc S.
;      REGISTRATION NUMBER:  32,181
;      REFERENCE/DOCKET NUMBER:  234-252P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 205-8000
;      TELEFAX:  (703) 205-8050
;      TELEX:  248345
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  468 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-204-656B-2

  Query Match             69.6%;  Score 32;  DB 1;  Length 468;
  Best Local Similarity   66.7%;  Pred. No. 1.1e-02;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 YDDQESVKS 9
              |||:  |:|
Db        156 YDDQAQVQS 164


RESULT 11
US-08-204-656B-4
; Sequence 4, Application US/08204656B
; Patent No. 5538882
;  GENERAL INFORMATION:
;    APPLICANT:  Matsui, Ikuo
;    APPLICANT:  Ishikawa, Kazuhiko
;    APPLICANT:  Miyairi, Sachio
;    APPLICANT:  Honda, Koichi
;    TITLE OF INVENTION:  Variant-Type Carbohydrate Hydrolase,
;    TITLE OF INVENTION:  Variant Gene Of The Enzyme And Method For Producing
;    TITLE OF INVENTION:  Oligosaccharide Using The Enzyme
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Birch, Stewart, Kolasch & Birch, LLP
;      STREET:  8110 Gatehouse Road, Suite 500 East
;      CITY:  Falls Church
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22042
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/204,656B
;      FILING DATE:  02-MAR-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Weiner, Marc S.
;      REGISTRATION NUMBER:  32,181
;      REFERENCE/DOCKET NUMBER:  234-252P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 205-8000
;      TELEFAX:  (703) 205-8050
;      TELEX:  248345
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  468 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-204-656B-4

  Query Match             69.6%;  Score 32;  DB 1;  Length 468;
  Best Local Similarity   66.7%;  Pred. No. 1.1e+02;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 YDDQESVKS 9
              |||:  |:|
Db        156 YDDQAQVQS 164


RESULT 12
US-08-204-656B-6
; Sequence 6, Application US/08204656B
; Patent No. 5538882
;  GENERAL INFORMATION:
;    APPLICANT:  Matsui, Ikuo
;    APPLICANT:  Ishikawa, Kazuhiko
```

```
;    APPLICANT:  Miyairi, Sachio
;    APPLICANT:  Honda, Koichi
;    TITLE OF INVENTION:  Variant-Type Carbohydrate Hydrolase,
;    TITLE OF INVENTION:  Variant Gene Of The Enzyme And Method For Producing
;    TITLE OF INVENTION:  Oligosaccharide Using The Enzyme
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Birch, Stewart, Kolasch & Birch, LLP
;      STREET:  8110 Gatehouse Road, Suite 500 East
;      CITY:  Falls Church
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22042
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/204,656B
;      FILING DATE:  02-MAR-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Weiner, Marc S.
;      REGISTRATION NUMBER:  32,181
;      REFERENCE/DOCKET NUMBER:  234-252P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 205-8000
;      TELEFAX:  (703) 205-8050
;      TELEX:  248345
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  468 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-204-656B-6

  Query Match             69.6%;  Score 32;  DB 1;  Length 468;
  Best Local Similarity   66.7%;  Pred. No. 1.1e+02;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 YDDQESVKS 9
              |||  |:|
Db        156 YDDQAQVQS 164


RESULT 13
US-08-204-656B-8
; Sequence 8, Application US/08204656B
; Patent No. 5538882
;  GENERAL INFORMATION:
;    APPLICANT:  Matsui, Ikuo
;    APPLICANT:  Ishikawa, Kazuhiko
;    APPLICANT:  Miyairi, Sachio
;    APPLICANT:  Honda, Koichi
;    TITLE OF INVENTION:  Variant-Type Carbohydrate Hydrolase,
;    TITLE OF INVENTION:  Variant Gene Of The Enzyme And Method For Producing
;    TITLE OF INVENTION:  Oligosaccharide Using The Enzyme
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Birch, Stewart, Kolasch & Birch, LLP
;      STREET:  8110 Gatehouse Road, Suite 500 East
;      CITY:  Falls Church
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22042
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/204,656B
;      FILING DATE:  02-MAR-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Weiner, Marc S.
;      REGISTRATION NUMBER:  32,181
;      REFERENCE/DOCKET NUMBER:  234-252P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 205-8000
;      TELEFAX:  (703) 205-8050
;      TELEX:  248345
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  468 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-204-656B-8

  Query Match             69.6%;  Score 32;  DB 1;  Length 468;
  Best Local Similarity   66.7%;  Pred. No. 1.1e+02;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 YDDQESVKS 9
              |||  |:|
Db        156 YDDQAQVQS 164


RESULT 14
US-08-470-702-6
; Sequence 6, Application US/08470702
; Patent No. 5631149
;  GENERAL INFORMATION:
;    APPLICANT:  MATSUI, IKUO
;    APPLICANT:  ISHIKAWA, KAZUHIKO
;    APPLICANT:  MIYAIRI, SACHIO
;    APPLICANT:  HONDA, KOICHI
;    TITLE OF INVENTION:  VARIANT-TYPE CARBOHYDRATE HYDROLASE,
;    TITLE OF INVENTION:  VARIANT GENE OF THE ENZYME AND METHOD FOR PRODUCING
;    TITLE OF INVENTION:  OLIGOSACCHARIDE USING THE ENZYME
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BIRCH, STEWART, KOLASCH & BIRCH
;      STREET:  8110 Gatehouse Road, Suite 500 East
;      CITY:  Falls Church
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22042
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/470,702
;      FILING DATE:  06-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/204,656
;      FILING DATE:  02-MAR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  WEINER, MARC S.
;      REGISTRATION NUMBER:  32,181
;      REFERENCE/DOCKET NUMBER:  234-252P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 205-8000
;      TELEFAX:  (703) 205-8050
;      TELEX:  248345
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  468 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
```

```
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-470-702-6

  Query Match             69.6%;  Score 32;  DB 1;  Length 468;
  Best Local Similarity   66.7%;  Pred. No. 1.1e-02;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 YDDQESVKS 9
              |:|| |:|
Db        156 YDDQAQVQS 164


RESULT 15
US-08-470-702-7
; Sequence 7, Application US/08470702
; Patent No. 5631149
;  GENERAL INFORMATION:
;    APPLICANT:  MATSUI, IKUO
;    APPLICANT:  ISHIKAWA, KAZUHIKO
;    APPLICANT:  MIYAIRI, SACHIO
;    APPLICANT:  HONDA, KOICHI
;    TITLE OF INVENTION:  VARIANT-TYPE CARBOHYDRATE HYDROLASE,
;    TITLE OF INVENTION:  VARIANT GENE OF THE ENZYME AND METHOD FOR PRODUCING
;    TITLE OF INVENTION:  OLIGOSACCHARIDE USING THE ENZYME
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BIRCH, STEWART, KOLASCH & BIRCH
;      STREET:  8110 Gatehouse Road, Suite 500 East
;      CITY:  Falls Church
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22042
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/470,702
;      FILING DATE:  06-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/204,656
;      FILING DATE:  02-MAR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  WEINER, MARC S.
;      REGISTRATION NUMBER:  32,181
;      REFERENCE/DOCKET NUMBER:  234-252P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 205-8000
;      TELEFAX:  (703) 205-8050
;      TELEX:  248345
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  468 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-470-702-7

  Query Match             69.6%;  Score 32;  DB 1;  Length 468;
  Best Local Similarity   66.7%;  Pred. No. 1.1e-02;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 YDDQESVKS 9
              |:|| |:|
Db        156 YDDQAQVQS 164


Search completed: September 24, 2003, 13:51:14
Job time : 11.9213 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 24, 2003, 13:47:40 ; Search time 20.1236 Seconds
(without alignments)
67.670 Million cell updates/sec

Title: US-09-744-282-3
Perfect score: 46
Sequence: 1 YDDQESVKS 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 566894 seqs, 151307093 residues

Total number of hits satisfying chosen parameters: 566894

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	46	100.0	383	15	US-10-097-340-45	Sequence 45, Appl
2	45	93.5	383	10	US-09-459-749D-17	Sequence 17, Appl
3	42	91.3	14	10	US-09-826-290-324	Sequence 324, App
4	42	91.3	14	11	US-09-791-393-109	Sequence 109, App
5	42	91.3	14	11	US-09-791-389-109	Sequence 109, App
6	33	71.7	85	10	US-09-522-334-28	Sequence 28, Appl
7	31	67.4	188	9	US-09-840-787-39	Sequence 39, Appl
8	31	67.4	188	11	US-09-883-152-8	Sequence 8, Appli
9	31	67.4	188	15	US-10-205-823-260	Sequence 260, App
10	31	67.4	188	15	US-10-177-293-288	Sequence 288, App
11	30	65.2	63	10	US-09-925-360-1042	Sequence 1042, Ap
12	30	65.2	131	9	US-09-864-761-35867	Sequence 35867, A
13	30	65.2	173	15	US-10-135-807-3	Sequence 3, Appli
14	30	65.2	202	15	US-10-156-761-11822	Sequence 11822, A
15	30	65.2	216	15	US-10-156-761-11988	Sequence 11988, A
16	30	65.2	346	11	US-09-847-208-67	Sequence 67, Appl
17	30	65.2	367	11	US-09-847-208-109	Sequence 109, App
18	30	65.2	449	9	US-09-815-242-12063	Sequence 12063, A
19	30	65.2	489	9	US-09-956-425-6	Sequence 6, Appli
20	30	65.2	526	15	US-10-187-267A-47	Sequence 47, Appl
21	30	65.2	853	12	US-10-259-165-260	Sequence 260, App
22	30	65.2	905	12	US-10-265-972-4	Sequence 4, Appli
23	30	65.2	969	9	US-09-118-276-2	Sequence 2, Appli
24	30	65.2	980	9	US-09-118-276-11	Sequence 11, Appl
25	30	65.2	1010	15	US-10-163-587A-6	Sequence 6, Appli
26	30	65.2	1014	12	US-10-369-378-25	Sequence 25, Appl
27	30	65.2	1014	15	US-10-097-340-4	Sequence 4, Appli
28	30	65.2	1014	15	US-10-163-587A-4	Sequence 4, Appli
29	30	65.2	1014	15	US-10-163-587A-18	Sequence 18, Appl
30	30	65.2	1014	15	US-10-163-587A-37	Sequence 37, Appl
31	30	65.2	1014	15	US-10-163-587A-38	Sequence 38, Appl
32	30	65.2	1014	15	US-10-163-587A-39	Sequence 39, Appl
33	30	65.2	1014	15	US-10-163-587A-40	Sequence 40, Appl
34	30	65.2	1063	12	US-10-369-378-47	Sequence 47, Appl
35	30	65.2	1217	12	US-10-311-406-2	Sequence 2, Appli
36	30	65.2	1324	12	US-10-205-219-113	Sequence 113, App
37	30	65.2	1531	9	US-09-876-889-347	Sequence 347, App
38	30	65.2	1531	10	US-09-998-598-2593	Sequence 2593, Ap
39	30	65.2	1531	12	US-10-301-822-211	Sequence 211, App
40	30	65.2	1531	15	US-10-171-311-222	Sequence 222, App
41	29	63.0	44	11	US-09-813-153-180	Sequence 180, App
42	29	63.0	174	15	US-10-304-928-28	Sequence 28, Appl
43	29	63.0	240	11	US-09-805-354-13	Sequence 13, Appl
44	29	63.0	240	15	US-10-144-259-13	Sequence 13, Appl
45	29	63.0	290	15	US-10-102-806-719	Sequence 719, App

ALIGNMENTS

RESULT 1
US-10-097-340-45
; Sequence 45, Application US/10097340
; Publication No. US2003C087250A1
; GENERAL INFORMATION:
; APPLICANT: John MONAHAN
; APPLICANT: Manjula GANNAVARAPU
; APPLICANT: Sebastian HOERSCH
; APPLICANT: Shubhangi KAMATKAR
; APPLICANT: Steve G. KOVATS
; APPLICANT: Rachel E. MEYERS
; APPLICANT: Michael MORRISEY
; APPLICANT: Peter OLANDT
; APPLICANT: Ami SEN
; APPLICANT: Peter VEIBY
; APPLICANT: Gordon B. MILLS
; APPLICANT: Robert C. BAST, Jr.
; APPLICANT: Karen LU
; APPLICANT: Rosemarie SCHMANDT
; APPLICANT: Xumei ZHAO
; APPLICANT: Karen GLATT
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
; TITLE OF INVENTION: Assessment, Prevention, and Therapy of Ovarian Cancer
; FILE REFERENCE: MR1-030
; CURRENT APPLICATION NUMBER: US/10/097,340
; CURRENT FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 60/276,025
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/325,149
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/276,026
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/324,967
; PRIOR FILING DATE: 2001/09/26
; PRIOR APPLICATION NUMBER: 60/311,732
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/325,102
; PRIOR FILING DATE: 2001-09-26

```
;  PRIOR APPLICATION NUMBER: 60/323,580
;  PRIOR FILING DATE: 2001-09-19
;  NUMBER OF SEQ ID NOS: 363
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 45
;   LENGTH: 383
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-097-340-45

  Query Match             100.0%;  Score 46;  DB 15;  Length 383;
  Best Local Similarity   100.0%;  Pred. No. 0.58;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YDDQESVKS 9
              |||||||||
Db        328 YDDQESVKS 336


RESULT 2
US-09-459-749D-17
; Sequence 17, Application US/09459749D
; Patent No. US20020136715A1
; GENERAL INFORMATION:
;  APPLICANT: Millis, Albert J. T.
;  TITLE OF INVENTION: Compositions and Methods For Altering Cell Migration
;  FILE REFERENCE: 0794.019A
;  CURRENT APPLICATION NUMBER: US/09/459,749D
;  CURRENT FILING DATE:  1999-12-10
;  PRIOR APPLICATION NUMBER: 60/111,856
;  PRIOR FILING DATE: 1998-12-11
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
;   LENGTH: 383
;   TYPE: PRT
;   ORGANISM: Sus scrofa
US-09-459-749D-17

  Query Match             93.5%;  Score 43;  DB 10;  Length 383;
  Best Local Similarity   88.9%;  Pred. No. 2.1;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YDDQESVKS 9
              |||||||||:
Db        328 YDDQESVKN 336


RESULT 3
US-09-826-290-324
; Sequence 324, Application US/09826290
; Patent No. US20020164668A1
; GENERAL INFORMATION:
;  APPLICANT: Durham, L.Kathryn
;  APPLICANT:  Friedman, David L.
;  APPLICANT:  Herath, Herath Mudiyanselage Athula Chandrasiri
;  APPLICANT:  Kimmel, Nida H.
;  APPLICANT:  Parekh, Rajesh Bhikhu
;  APPLICANT:  Potter, David M.
;  APPLICANT:  Rohlff, Christian
;  APPLICANT:  Silber, B. Michael
;  APPLICANT:  Stiger, Thomas R.
;  APPLICANT:  Sunderland, P. Trey
;  APPLICANT:  Townsend, Robert Reid
;  APPLICANT:  White, Frost
;  APPLICANT:  Williams, Stephen A.
;  TITLE OF INVENTION: Nucleic Acid Molecules, Polypeptides and
;  TITLE OF INVENTION:  Uses Therefor, Including Diagnosis and Treatment of
;  TITLE OF INVENTION:  Alzheimer's Disease
;  FILE REFERENCE: 2572-1-001 N2
;  CURRENT APPLICATION NUMBER: US/09/826,290
;  CURRENT FILING DATE:  2001-04-30
   PRIOR APPLICATION NUMBER: US 60/194,504
;  PRIOR FILING DATE: 2000-04-03
;  PRIOR APPLICATION NUMBER: US 60/253,647
;  PRIOR FILING DATE: 2000-11-28
;  NUMBER OF SEQ ID NOS: 492
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 324
;   LENGTH: 14
;   TYPE: PRT
;   ORGANISM: homo sapien
US-09-826-290-324

  Query Match             91.3%;  Score 42;  DB 10;  Length 14;
  Best Local Similarity   100.0%;  Pred. No. 0.097;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YDDQESVK 8
              ||||||||
Db          7 YDDQESVK 14


RESULT 4
US-09-791-393-109
; Sequence 109, Application US/09791393
; Publication No. US20030032200A1
; GENERAL INFORMATION:
;  APPLICANT: Herath, Herath Mudiyanselage Athula Chandrasiri
;  APPLICANT: Parekh, Rajesh Bhikhu
;  APPLICANT: Rohlff, Christian
;  TITLE OF INVENTION: Proteins, Genes and Their Use for
;  TITLE OF INVENTION: Diagnosis and Treatment of Bipolar Affective Disorder (BAD)
;  TITLE OF INVENTION: and Unipolar Depression
;  FILE REFERENCE: 2543-1-001 N1
;  CURRENT APPLICATION NUMBER: US/09/791,393
;  CURRENT FILING DATE: 2002-01-02
;  EARLIER APPLICATION NUMBER: GB 0004412.3
;  EARLIER FILING DATE: 2000-02-24
;  EARLIER APPLICATION NUMBER: GB 0030050.9
;  EARLIER FILING DATE: 2000-12-08
;  EARLIER APPLICATION NUMBER: US 60/254,830
;  EARLIER FILING DATE: 2000-12-12
;  NUMBER OF SEQ ID NOS: 308
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 109
;   LENGTH: 14
;   TYPE: PRT
;   ORGANISM: homo sapien
US-09-791-393-109

  Query Match             91.3%;  Score 42;  DB 11;  Length 14;
  Best Local Similarity   100.0%;  Pred. No. 0.097;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YDDQESVK 8
              ||||||||
Db          7 YDDQESVK 14


RESULT 5
US-09-791-389-109
; Sequence 109, Application US/09791389
; Publication No. US20030032773A1
; GENERAL INFORMATION:
;  APPLICANT: Herath, Herath Mudiyanselage Athula Chandrasiri
;  APPLICANT:  Parekh, Rajesh Bhikhu
;  APPLICANT:  Rohlff, Christian
;  APPLICANT:  Terrett, Jonathan Alexander
;  APPLICANT:  Tyson, Kerry Louise
;  TITLE OF INVENTION: Proteins, Genes and Their Use for
;  TITLE OF INVENTION:  Diagnosis and Treatment of Bipolar Affective Disorder (BAD)
;  TITLE OF INVENTION:  and Unipolar Depression
;  FILE REFERENCE: 2543-1-001 N2
;  CURRENT APPLICATION NUMBER: US/09/791,389
;  CURRENT FILING DATE:  2001-02-23
```

```
;  PRIOR APPLICATION NUMBER: GB 0004412.3
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: GB 0030050.9
;  PRIOR FILING DATE: 2000-12-08
;  PRIOR APPLICATION NUMBER: US 60/254,830
;  PRIOR FILING DATE: 2000-12-12
;  NUMBER OF SEQ ID NOS: 308
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 109
;   LENGTH: 14
;   TYPE: PRT
;   ORGANISM: homo sapien
US-09-791-389-109

  Query Match             91.3%;  Score 42;  DB 11;  Length 14;
  Best Local Similarity   100.0%;  Pred. No. 0.097;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YDDQESVK 8
              ||||||||
Db          7 YDDQESVK 14


RESULT 6
US-09-522-334-28
; Sequence 28, Application US/09522334
; Patent No. US20020157130A1
; GENERAL INFORMATION:
;  APPLICANT:  Wagner, Ry
;  APPLICANT:  Mathews, Helena
;  APPLICANT:  Liu, Xing Liang
;  APPLICANT:  Waggoner, Wency J.
;  TITLE OF INVENTION:  TRAIT-ASSOCIATED GENE IDENTIFICATION
;  TITLE OF INVENTION:  METHOD
;  FILE REFERENCE:  4257-0018.3C
;  CURRENT APPLICATION NUMBER: US/09/522,334
;  CURRENT FILING DATE:  2000-03-09
;  PRIOR APPLICATION NUMBER:  US 60/124,232
;  PRIOR FILING DATE:  1999-03-12
;  NUMBER OF SEQ ID NOS:  28
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 28
;   LENGTH: 85
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: ORF from plasmid rescue of L23
US-09-522-334-28

  Query Match             71.7%;  Score 33;  DB 10;  Length 85;
  Best Local Similarity   55.6%;  Pred. No. 34;
  Matches    5;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 YDDQESVKS 9
              |:|:| : |
Db         46 YDDEEKISS 54


RESULT 7
US-09-840-787-39
; Sequence 39, Application US/09840787
; Patent No. US20020058264A1
;   GENERAL INFORMATION:
;        APPLICANT: Lal, Preeti
;                   Hillman, Jennifer L.
;                   Bandman, Olga
;                   Shah, Purvi
;                   Au-Young, Janice
;                   Yue, Henry
;                   Guegler, Karl J.
;                   Corley, Neil C.
;        TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
;        NUMBER OF SEQUENCES: 98
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Incyte Pharmaceuticals, Inc.
;             STREET: 3174 Porter Drive
;             CITY: Palo Alto
;             STATE: CA
;             COUNTRY: USA
;             ZIP: 94304
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette
;             COMPUTER: IBM Compatible
;             OPERATING SYSTEM: DOS
;             SOFTWARE: FastSEQ for Windows Version 2.0
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/840,787
;             FILING DATE: 23-Apr-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/518,865
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Billings, Lucy J.
;             REGISTRATION NUMBER: 36,749
;             REFERENCE/DOCKET NUMBER: PF-0356 US
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 415-855-0555
;             TELEFAX: 415-845-4166
;             TELEX: <Unknown>
;   INFORMATION FOR SEQ ID NO: 39:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 188 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        IMMEDIATE SOURCE:
;             LIBRARY: BRAITUT03
;             CLONE: 2112230
;        SEQUENCE DESCRIPTION: SEQ ID NO: 39 :
US-09-840-787-39

  Query Match             67.4%;  Score 31;  DB 9;  Length 188;
  Best Local Similarity   75.0%;  Pred. No. 1.9e+02;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 DDQESVKS 9
              |:|| |||
Db         95 DEQEGVKS 102


RESULT 8
US-09-883-152-8
; Sequence 8, Application US/09883152
; Publication No. US20030008284A1
; GENERAL INFORMATION:
;  APPLICANT: Kennedy, Giulia
;  APPLICANT:  Kang, Sanmao
;  APPLICANT:  Reinhard, Christoph
;  APPLICANT:  Jefferson, Anne Bennett
;  TITLE OF INVENTION: POLYNUCLEOTIDES RELATED TO COLON CANCER
;  FILE REFERENCE: 2300-1663
;  CURRENT APPLICATION NUMBER: US/09/883,152
;  CURRENT FILING DATE:  2001-06-15
;  PRIOR APPLICATION NUMBER: 60/211,835
;  PRIOR FILING DATE: 2000-06-15
;  NUMBER OF SEQ ID NOS: 127
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8
;   LENGTH: 188
;   TYPE: PRT
;   ORGANISM: H. sapiens
US-09-883-152-8

  Query Match             67.4%;  Score 31;  DB 11;  Length 188;
  Best Local Similarity   75.0%;  Pred. No. 1.9e+02;
```

```
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 DDQESVKS 9
              |:|| |||
Db         95 DEQEGVKS 102


RESULT 9
US-10-205-823-260
; Sequence 260, Application US/10205823
; Publication No. US20030108963A1
; GENERAL INFORMATION:
;  APPLICANT: Schlegel, Robert
;  APPLICANT:  Monahan, John E.
;  APPLICANT:  Endege, Wilson O.
;  APPLICANT:  Gannavarapu, Manjula
;  APPLICANT:  Gorbatcheva, Bella
;  APPLICANT:  Hoersch, Sebastian
;  APPLICANT:  Kamatkar, Shubhangi
;  APPLICANT:  Wonsey, Angela M.
;  APPLICANT:  Glatt, Karen
;  APPLICANT:  Zhao, Xumei
;  APPLICANT:  Anderson, Dustin
;  TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
;  TITLE OF INVENTION:  METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
;  TITLE OF INVENTION:  THERAPY OF PROSTATE CANCER
;  FILE REFERENCE: MRI-044
;  CURRENT APPLICATION NUMBER: US/10/205,823
;  CURRENT FILING DATE:  2002-07-25
;  PRIOR APPLICATION NUMBER: 60/307,982
;  PRIOR FILING DATE: 2001-07-25
;  PRIOR APPLICATION NUMBER: 60/314,356
;  PRIOR FILING DATE: 2001-08-22
;  PRIOR APPLICATION NUMBER: 60/325,020
;  PRIOR FILING DATE: 2001-09-25
;  PRIOR APPLICATION NUMBER: 60/341,746
;  PRIOR FILING DATE: 2001-12-12
;  PRIOR APPLICATION NUMBER: 60/362,158
;  PRIOR FILING DATE: 2002-03-05
;  NUMBER OF SEQ ID NOS: 455
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 260
;   LENGTH: 188
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-205-823-260

  Query Match             67.4%;  Score 31;  DB 15;  Length 188;
  Best Local Similarity   75.0%;  Pred. No. 1.9e+02;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 DDQESVKS 9
              |:|| |||
Db         95 DEQEGVKS 102


RESULT 10
US-10-177-293-288
; Sequence 288, Application US/10177293
; Publication No. US20030124128A1
; GENERAL INFORMATION:
;  APPLICANT: Lillie, James
;  APPLICANT:  Glatt, Karen
;  APPLICANT:  Zhao, Xumei
;  APPLICANT:  Gannavarpu, Manjula
;  APPLICANT:  Kamatkar, Shubhangi
;  APPLICANT:  Mertens, Maureen
;  APPLICANT:  Myer, Vic
;  APPLICANT:  Wang, Youzhen
;  APPLICANT:  Xu, Yongyao
;  APPLICANT:  Hoersch, Sebastian
;  APPLICANT:  Monahan, John
;  APPLICANT:  Meyers, Rachel E.
;  APPLICANT:  Bast Jr., Robert C.
;  APPLICANT:  Hortobagyi, Gabriel N.
;  APPLICANT:  Pusztai, Lajos
;  APPLICANT:  Meric, Funda
;  APPLICANT:  Sahin, Aysegul
;  APPLICANT:  Mills, Gordon B.
;  TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMEN
;  TITLE OF INVENTION:  PREVENTION, AND THERAPY OF BREAST CANCER
;  FILE REFERENCE: MRI-038
;  CURRENT APPLICATION NUMBER: US/10/177,293
;  CURRENT FILING DATE:  2002-06-21
;  PRIOR APPLICATION NUMBER: US 60/299,887
;  PRIOR FILING DATE: 2001-06-21
;  PRIOR APPLICATION NUMBER: US 60/301,572
;  PRIOR FILING DATE: 2001-06-27
;  PRIOR APPLICATION NUMBER: US 60/306,501
;  PRIOR FILING DATE: 2001-07-18
;  PRIOR APPLICATION NUMBER: US 60/325,002
;  PRIOR FILING DATE: 2001-09-25
;  PRIOR APPLICATION NUMBER: US 60/362,585
;  PRIOR FILING DATE: 2002-03-05
;  PRIOR APPLICATION NUMBER: US 60/xxx,xxx
;  PRIOR FILING DATE: 2002-05-14
;  NUMBER OF SEQ ID NOS: 506
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 288
;   LENGTH: 188
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-177-293-288

  Query Match             67.4%;  Score 31;  DB 15;  Length 188;
  Best Local Similarity   75.0%;  Pred. No. 1.9e+02;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 DDQESVKS 9
              |:|| |||
Db         95 DEQEGVKS 102


RESULT 11
US-09-925-300-1042
; Sequence 1042, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
;  APPLICANT: Craig Rosen,
;  APPLICANT:  Steve Ruben
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA101
;  CURRENT APPLICATION NUMBER: US/09/925,300
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05988
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 1890
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1042
;   LENGTH: 63
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (14)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-300-1042

  Query Match             65.2%;  Score 30;  DB 10;  Length 63;
  Best Local Similarity   71.4%;  Pred. No. 93;
  Matches    5;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YDDQESV 7
              :|||:||
```

```
Db          22 FDDEESV 28


RESULT 12
US-09-864-761-35867
; Sequence 35867, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 35867
;   LENGTH: 131
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AP000516.1
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.5
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.2
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.92
;   OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.3
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL =  1
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.1
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.5
;   OTHER INFORMATION: SWISSPROT HIT: P14373, EVALUE 1.00e-24
;   OTHER INFORMATION: EST_HUMAN HIT: BE293328.1, EVALUE 2.00e-51
US-09-864-761-35867

  Query Match             65.2%;  Score 30;  DB 9;  Length 131;
  Best Local Similarity   71.4%;  Pred. No. 2e-02;
  Matches    5;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YDDQESV 7
              |:|||:|
Db         83 YEDQEAV 89


RESULT 13
US-10-135-807-3
; Sequence 3, Application US/10135807
; Publication No. US20030049655A1
; GENERAL INFORMATION:
;  APPLICANT: The Secretary of State for Defence in Her Britannic Majesty's
;  APPLICANT:  Government of the United Kingdom of Great Britain and
;  APPLICANT:  No. US20030049655Althern Ireland
;  APPLICANT:  Clark, Duncan Roy
;  APPLICANT:  Vincent, Suzanne P
;  TITLE OF INVENTION: Amplification process
;  FILE REFERENCE: CG/P/133/WOD
;  CURRENT APPLICATION NUMBER: US/10/135,807
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: GB 0110501.4
;  PRIOR FILING DATE: 2001-04-30
;  NUMBER OF SEQ ID NOS: 26
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
;   LENGTH: 173
;   TYPE: PRT
;   ORGANISM: Sulfolobus solfataricus
US-10-135-807-3

  Query Match             65.2%;  Score 30;  DB 15;  Length 173;
  Best Local Similarity   71.4%;  Pred. No. 2.7e+02;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 YDDQESV 7
              |||:| |
Db         29 YDDEEGV 35


RESULT 14
US-10-156-761-11822
; Sequence 11822, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
;  APPLICANT: OMURA, SATOSHI
;  APPLICANT:  IKEDA, HARUO
;  APPLICANT:  ISHIKAWA, JUN
;  APPLICANT:  HORIKAWA, HIROSHI
;  APPLICANT:  SHIBA, TADAYOSHI
;  APPLICANT:  SAKAKI, YOSHIYUKI
;  APPLICANT:  HATTORI, MASAHIRA
;  TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;  FILE REFERENCE: 249-262
;  CURRENT APPLICATION NUMBER: US/10/156,761
;  CURRENT FILING DATE:  2002-05-29
;  PRIOR APPLICATION NUMBER: JP 2001-204089
;  PRIOR FILING DATE: 2001-05-30
;  PRIOR APPLICATION NUMBER: JP 2001-272697
;  PRIOR FILING DATE: 2001-08-02
;  NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 11822
;   LENGTH: 202
;   TYPE: PRT
;   ORGANISM: Streptomyces avermitilis
US-10-156-761-11822

  Query Match             65.2%;  Score 30;  DB 15;  Length 202;
  Best Local Similarity   62.5%;  Pred. No. 3.2e+02;
  Matches    5;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 YDDQESVK 8
```

```
              |:|:| ||
Db         88 YEDREGVK 95


RESULT 15
US-10-156-761-11988
; Sequence 11988, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
;  APPLICANT: OMURA, SATOSHI
;  APPLICANT:  IKEDA, HARUO
;  APPLICANT:  ISHIKAWA, JUN
;  APPLICANT:  HORIKAWA, HIROSHI
;  APPLICANT:  SHIBA, TADAYOSHI
;  APPLICANT:  SAKAKI, YOSHIYUKI
;  APPLICANT:  HATTORI, MASAHIRA
;  TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;  FILE REFERENCE: 249-262
;  CURRENT APPLICATION NUMBER: US/10/156,761
;  CURRENT FILING DATE:  2002-05-29
;  PRIOR APPLICATION NUMBER: JP 2001-204089
;  PRIOR FILING DATE: 2001-05-30
;  PRIOR APPLICATION NUMBER: JP 2001-272697
;  PRIOR FILING DATE: 2001-08-02
;  NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 11988
;   LENGTH: 216
;   TYPE: PRT
;   ORGANISM: Streptomyces avermitilis
US-10-156-761-11988

  Query Match             65.2%;  Score 30;  DB 15;  Length 216;
  Best Local Similarity   75.0%;  Pred. No. 3.5e-02;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 DDQESVKS 9
              :|:|| |:|
Db          9 DDQELVRS 16


Search completed: September 24, 2003, 14:22:27
Job time : 21.1236 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 24, 2003, 13:34:08 ; Search time 10.9213 Seconds
(without alignments)
79.250 Million cell updates/sec

Title: US-09-744-282-3
Perfect score: 46
Sequence: 1 YDDQESVKS 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	46	100.0	383	2	A49562	cartilage glycopro
2	43	93.5	383	2	S51327	heparin-binding gl
3	37	80.4	4063	2	T42993	probable spectrin
4	37	80.4	4101	2	T23630	hypothetical prote
5	36	78.3	797	2	S28103	probable DNA-direc
6	36	78.3	5170	2	T15348	hypothetical prote
7	35	76.1	124	2	A45691	hypothetical prote
8	35	76.1	235	2	C85115	cytochrome c-type
9	35	76.1	235	2	G97984	cytochrome c-type
10	35	76.1	405	2	S61551	breast-regressing
11	34	73.9	441	2	S60618	4-alpha-glucanotra
12	34	73.9	1612	2	S59969	DNA topoisomerase
13	34	73.9	1626	2	A39242	DNA topoisomerase
14	33	71.7	374	2	C81674	hemolysin, probabl
15	33	71.7	546	1	F64835	paraquat-inducible
16	33	71.7	546	2	C90758	paraquat-inducible
17	33	71.7	546	2	A85622	paraquat-inducible
18	33	71.7	546	2	AC0626	probable secreted
19	33	71.7	827	2	S50714	DNA-binding protei
20	33	71.7	1058	2	T06935	CCP1-interacting p
21	33	71.7	1199	2	JC4815	major surface glyc
22	33	71.7	1248	2	I05059	hypothetical prote
23	32	69.6	123	2	F45691	hypothetical prote
24	32	69.6	133	2	AI0098	hypothetical prote
25	32	69.6	176	2	E72348	conserved hypothet
26	32	69.6	230	2	T31722	hypothetical prote
27	32	69.6	369	2	JC4292	probable transposa
28	32	69.6	373	2	S15161	long-chain-fatty-a
29	32	69.6	413	2	T43810	methylaspartate am
30	32	69.6	413	2	A39724	3-methylaspartate
31	32	69.6	413	2	A85575	probable methylasp
32	32	69.6	439	2	H81414	probable heat shoc
33	32	69.6	494	1	ALBYAP	alpha-amylase (EC
34	32	69.6	516	2	T00974	probable SF16 prot
35	32	69.6	518	2	T15404	hypothetical prote
36	32	69.6	609	2	T28896	hypothetical prote
37	32	69.6	655	2	A48468	dnaK-type molecula
38	32	69.6	689	2	T40364	hypothetical prote
39	32	69.6	771	2	G71409	probable replicati
40	32	69.6	773	2	S46011	probable pre-mRNA-
41	32	69.6	875	2	S75377	probable helicase
42	32	69.6	2946	2	T00867	hypothetical prote
43	31.5	68.5	569	2	T22372	hypothetical prote
44	31	67.4	118	2	C37927	hypothetical prote
45	31	67.4	118	2	B95058	transcription regu

ALIGNMENTS

RESULT 1
A49562
cartilage glycoprotein gp39 precursor - human
N;Alternate names: 39K synovial protein
C;Species: Homo sapiens (man)
C;Date: 23-Mar-1995 #sequence_revision 23-Mar-1995 #text_change 22-Jun-1999
C;Accession: A49562; S10677; A33162
R;Hakala, B.E.; White, C.; Recklies, A.D.
J. Biol. Chem. 268, 25803-25810, 1993
A;Title: Human cartilage gp-39, a major secretory product of articular chondrocytes
A;Reference number: A49562; MUID:94064658; PMID:8245017
A;Accession: A49562
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-383 <HAK>
A;Cross-references: GB:M80927; NID:g348911; PIDN:AAA16074.1; PID:g348912
R;Nyirkos, P.; Golds, E.E.
Biochem. J. 269, 265-268, 1990
A;Title: Human synovial cells secrete a 39 kDa protein similar to a bovine mammary
A;Reference number: S10677; MUID:90328983; PMID:2375755
A;Accession: S10677
A;Molecule type: protein
A;Residues: 22-40,'X',42-45 <NY2>
C;Superfamily: Streptomyces chitinase ch140
C;Keywords: cartilage; extracellular protein; glycoprotein
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-383/Product: cartilage glycoprotein gp39 #status predicted <MAT>

Query Match 100.0%; Score 46; DB 2; Length 383;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 YDDQESVKS 9
              |||||||||
Db        328 YDDQESVKS 336
```

RESULT 2
S51327
heparin-binding glycoprotein 38K - pig
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 22-Jun-1999
C;Accession: S51327
R;Shackelton, L.M.; Mann, D.M.; Millis, A.J.T.
submitted to the EMBL Data Library, January 1995
A;Description: Identification of a 38kDa heparin-binding glycoprotein (gp38k) in di
A;Reference number: S51327
A;Accession: S51327
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-383 <SHA>
A;Cross-references: EMBL:Z47803; NID:g634097; PIDN:CAA87764.1; PID:g634098

```
C;Superfamily: Streptomyces chitinase ch140

  Query Match            93.5%;  Score 43;  DB 2;  Length 383;
  Best Local Similarity  88.9%;  Pred. No. 0.64;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YDDQESVKS 9
              |||||||||
Db        328 YDDQESVKN 336


RESULT 3
T42993
probable spectrin beta chain - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 11-Jan-2000
C;Accession: T42993
R;Austin, J.; Praitis, V.; McKeown, C.
submitted to the EMBL Data Library, March 1998
A;Description: Sma-1 encodes a BH-spectrin homolog required for C. elegans morphogenesis
A;Reference number: Z22279
A;Accession: T42993
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-4063 <AUS>
A;Cross-references: EMBL:AF053496; PIDN:AAC08577.1
C;Genetics:
A;Gene: sma1

  Query Match            80.4%;  Score 37;  DB 2;  Length 4063;
  Best Local Similarity  66.7%;  Pred. No. 1.4e-02;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 YDDQESVKS 9
              ||| :|:||
Db       4036 YDDSDSIKS 4044


RESULT 4
T23630
hypothetical protein R31.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999
C;Accession: T23630; T24242
R;Kershaw, J.
submitted to the EMBL Data Library, November 1996
A;Reference number: Z19774
A;Accession: T23630
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-4101 <WIL>
A;Cross-references: EMBL:Z81570; PIDN:CAB04608.1; GSPDB:GN00023; CESP:R31.1
A;Experimental source: clone K12G11
R;Lennard, N.
submitted to the EMBL Data Library, July 1996
A;Reference number: Z19852
A;Accession: T24242
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-4101 <WI2>
A;Cross-references: EMBL:Z75956; PIDN:CAB00130.1; GSPDB:GN00023; CESP:R31.1
A;Experimental source: clone R31
C;Genetics:
A;Gene: CESP:R31.1
A;Map position: 5
A;Introns: 29/3; 89/3; 184/3; 312/3; 812/2; 1461/3; 1531/3; 2050/2; 2413/2; 2466/2; 3008

  Query Match            80.4%;  Score 37;  DB 2;  Length 4101;
  Best Local Similarity  66.7%;  Pred. No. 1.4e-02;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 YDDQESVKS 9
              ||| :|:||
Db       4058 YDDSDSIKS 4066


RESULT 5
S28103
probable DNA-directed DNA polymerase (EC 2.7.7.7) - gill mushroom (Agaricus bitorqui
C;Species: Agaricus bitorquis
C;Date: 28-May-1993 #sequence_revision 28-May-1993 #text_change 23-Feb-1997
C;Accession: S28103
R;Robison, M.M.; Royer, J.C.; Horgen, P.A.
Curr. Genet. 19, 495-502, 1991
A;Title: Homology between mitochondrial DNA of Agaricus bisporus and an internal por
A;Reference number: S28103; MUID:91347410; PMID:1879001
A;Accession: S28103
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-797 <ROB>
A;Cross-references: EMBL:Z11513
C;Genetics:
A;Genome: plasmid
C;Keywords: DNA binding; mitochondrion; nucleotidyltransferase

  Query Match            78.3%;  Score 36;  DB 2;  Length 797;
  Best Local Similarity  75.0%;  Pred. No. 38;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 YDDQESVK 8
              |||| |:|
Db        400 YDDQNSIK 407


RESULT 6
T15348
hypothetical protein B0350.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C;Accession: T15348
R;Gatting, S.
submitted to the EMBL Data Library, February 1996
A;Description: The sequence of C. elegans cosmid B0350.
A;Reference number: Z18332
A;Accession: T15348
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-5170 <GAT>
A;Cross-references: EMBL:U50071; NID:g1208871; PID:g1208877; PIDN:AAA93447.1; CESP:B
C;Genetics:
A;Gene: CESP:B0350.1
A;Introns: 48/1; 5039/3; 5116/3

  Query Match            78.3%;  Score 36;  DB 2;  Length 5170;
  Best Local Similarity  87.5%;  Pred. No. 2.9e+02;
  Matches    7;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 DDQESVKS 9
              |||||:||
Db       2119 DDQESLKS 2126


RESULT 7
A45691
hypothetical protein 1 [imported] - Lactobacillus delbrueckii subsp. lactis phage LL
C;Species: Lactobacillus delbrueckii subsp. lactis phage LL-H
C;Date: 24-Feb-1994 #sequence_revision 25-Apr-1997 #text_change 28-Jul-2000
C;Accession: A45691
R;Vasala, A.; Dupont, L.; Baumann, M.; Ritzenthaler, P.; Alatossava, T.
J. Virol. 67, 3061-3068, 1993
A;Title: Molecular comparison of the structural proteins encoding gene clusters of t
A;Reference number: A45691; MUID:93267750; PMID:8497043
A;Accession: A45691
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-124 <VAS>
```

A;Cross-references: GB:L02496; NID:g215199; PIDN:AAC00058.1; PID:g623066
A;Note: sequence extracted from NCBI backbone (NCBIN:132369, NCBIP:132365)

Query Match 76.1%; Score 35; DB 2; Length 124;
Best Local Similarity 55.6%; Pred. No. 8;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YDDQESVKS 9
Db 46 YDECFAIKN 54

RESULT 8
C95115
cytochrome c-type biogenesis protein CcdA [imported] - Streptococcus pneumoniae (strain
C;Species: Streptococcus pneumoniae
C;Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 24-Aug-2001
C;Accession: C95115
R;Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple,
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A;Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A;Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A;Reference number: A95000; MUID:21357209; PMID:11463916
A;Accession: C95115
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-235 <KUR>
A;Cross-references: GB:AE005672; PIDN:AAK75116.1; PID:g14972472; GSPDB:GN00164; TIGR:SP4
A;Experimental source: strain TIGR4
C;Genetics:
A;Gene: SP0999
C;Superfamily: Haemophilus influenzae probable holocytochrome-c synthase

Query Match 76.1%; Score 35; DB 2; Length 235;
Best Local Similarity 87.5%; Pred. No. 16;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 DDQESVKS 9
Db 36 DDQESAKS 43

RESULT 9
G97984
cytochrome c-type biogenesis protein (imported) - Streptococcus pneumoniae (strain R6)
C;Species: Streptococcus pneumoniae
C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 02-Nov-2001
C;Accession: G97984
R;Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszczak, L.; Burgett, S.; DeHoff, B.S.; E
e, R.; LeBlanc, D.J.; Lee, L.N.; Lofkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A;Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Caskunas, S.R.;
A;Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A;Reference number: A97872; MUID:21429245; PMID:11544234
A;Accession: G97984
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-235 <KUR>
A;Cross-references: GB:AE007317; PIDN:AAK99707.1; PID:g15458509; GSPDB:GN00174
C;Genetics:
A;Gene: ccdA
C;Superfamily: Haemophilus influenzae probable holocytochrome-c synthase

Query Match 76.1%; Score 35; DB 2; Length 235;
Best Local Similarity 87.5%; Pred. No. 16;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 DDQESVKS 9
Db 36 DDQESAKS 43

RESULT 10
S61551
breast-regressing protein brp39 precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
C;Accession: S61551; S61550; I48271
R;Morrison, B.W.; Leder, P.
Oncogene 9, 3417-3426, 1994
A;Title: neu and ras initiate murine mammary tumors that share genetic markers gene
A;Reference number: I48271; MUID:95060797; PMID:7970700
A;Accession: S61551
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-405 <MOR1>
A;Cross-references: EMBL:X93035; NID:g1085065; PIDN:CAA63603.1; PID:g1085066
R;Morrison, B.W.
submitted to the EMBL Data Library, November 1995
A;Reference number: S61550
A;Accession: S61550
A;Molecule type: mRNA
A;Residues: 1-245,'I',247-330,'H',332-350,'MVWALDLDDPQGTCQPKEFFPLTNAIKDALA' <MOR2>
A;Cross-references: EMBL:X93035; NID:g1085065; PIDN:CAA63603.1; PID:g1085066
A;Note: the differences at the carboxyl end are due to a frameshift error
C;Genetics:
A;Gene: brp39
C;Superfamily: Streptomyces chitinase chi40
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-405/Product: breast-regressing protein brp39 #status predicted <MAT>

Query Match 76.1%; Score 35; DB 2; Length 405;
Best Local Similarity 66.7%; Pred. No. 29;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YDDQESVKS 9
Db 329 YEDKESVKN 337

RESULT 11
S60618
4-alpha-glucanotransferase (EC 2.4.1.25) - Thermotoga maritima (strain MSB8)
C;Species: Thermotoga maritima
C;Date: 27-Apr-1996 #sequence_revision 08-Nov-1996 #text_change 21-Jul-2000
C;Accession: S60618; S23198; D72387
R;Heinrich, P.; Huber, W.; Liebl, W.
Syst. Appl. Microbiol. 17, 297-305, 1994
A;Title: Expression in Escherichia coli and structure of the gene encoding 4-alpha-
related enzyme subfamilies.
A;Reference number: S60618
A;Accession: S60618
A;Molecule type: DNA
A;Residues: 1-441 <HEI>
A;Cross-references: EMBL:Z50813; NID:g951310; PIDN:CAA90693.1; PID:g951311
R;Liebl, W.; Feil, R.; Gabelsberger, J.; Kellermann, J.; Schleifer, K.H.
Eur. J. Biochem. 207, 81-88, 1992
A;Title: Purification and characterization of a novel thermostable 4-alpha-glucanot
A;Reference number: S23198; MUID:92331687; PMID:1628664
A;Accession: S23198
A;Molecule type: protein
A;Residues: 'A',2-23 <LIE>
R;Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; H
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson
C.M.
Nature 399, 323-329, 1999
A;Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome
A;Reference number: A72200; MUID:99287316; PMID:10360571
A;Accession: D72387
A;Molecule type: DNA
A;Residues: 1-264,'G',266-441 <ARN>
A;Cross-references: GB:AE001716; GB:AE000512; NID:g4980853; PIDN:AAD35451.1; PID:g4
A;Experimental source: strain MSB8

```
C;Genetics:
A;Gene: TM0354
C;Keywords: glycosyltransferase; hexosyltransferase

  Query Match             73.9%;  Score 34;  DB 2;  Length 441;
  Best Local Similarity   75.0%;  Pred. No. 50;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 YDDQESVK 8
              |||| |:|
Db        409 YDDQHSLK 416


RESULT 12
S59969
DNA topoisomerase (ATP-hydrolyzing) (EC 5.99.1.3) isoform beta - Chinese hamster
N;Alternate names: DNA topoisomerase II isoform beta; DNA-gyrase
C;Species: Cricetulus griseus (Chinese hamster)
C;Date: 15-Feb-1996 #sequence_revision 13-Mar-1997 #text_change 16-Jul-1999
C;Accession: S59969; S54154
R;Dereuddre, S.; Frey, S.; Delaporte, C.; Jacquemin-Sablon, A.
Biochim. Biophys. Acta 1264, 178-182, 1995
A;Title: Cloning and characterization of full-length cDNAs coding for the DNA topoisomer
A;Reference number: S59969; MUID:96085121; PMID:7495861
A;Accession: S59969
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-1612 <DER>
A;Cross-references: EMBL:X86455; NID:g790987; PIDN:CAA60173.1; PID:g790988
A;Experimental source: lung
C;Superfamily: eukaryotic type II DNA topoisomerase; phage T4 DNA topoisomerase (ATP-hyd
C;Keywords: ATP; DNA binding; DNA replication; isomerase; nucleus
F;697-927/Domain: phage T4 DNA topoisomerase (ATP-hydrolyzing) medium chain homology <T4

  Query Match             73.9%;  Score 34;  DB 2;  Length 1612;
  Best Local Similarity   66.7%;  Pred. No. 2.1e-02;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 YDDQESVKS 9
              :|| |:|:|
Db        531 YDDAESLKT 539


RESULT 13
A39242
DNA topoisomerase (ATP-hydrolyzing) (EC 5.99.1.3) beta, splice form 2 - human
N;Alternate names: DNA topoisomerase II isoform beta-2
N;Contains: DNA topoisomerase II isoform beta-1
C;Species: Homo sapiens (man)
C;Date: 04-Oct-1991 #sequence_revision 03-May-1996 #text_change 21-Jul-2000
C;Accession: S26730; A39242; S10710; S33970; S30191; S41641; S30190
R;Jenkins, J.R.; Ayton, P.; Jones, T.; Davies, S.L.; Simmons, D.L.; Harris, A.L.; Sheer,
Nucleic Acids Res. 20, 5587-5592, 1992
A;Title: Isolation of cDNA clones encoding the beta isozyme of human DNA topoisomerase I
A;Reference number: S26730; MUID:93087165; PMID:1333583
A;Accession: S26730
A;Molecule type: mRNA
A;Residues: 1-23,29-1626 <JEN>
A;Cross-references: EMBL:X58060; NID:g37230; PIDN:CAA48197.1; PID:g37231
R;Chung, I.D.Y.; Drake, F.H.; Tan, K.B.; Per, S.R.; Crooke, S.T.; Mirabelli, C.K.
Proc. Natl. Acad. Sci. U.S.A. 86, 9431-9435, 1989
A;Title: Characterization and immunological identification of cDNA clones encoding two i
A;Reference number: A39242; MUID:90083281; PMID:2556712
A;Accession: A39242
A;Molecule type: mRNA
A;Residues: 149-1043 <CHU>
A;Cross-references: GB:M27504
R;Austin, C.A.; Fisher, L.M.
FEBS Lett. 266, 115-117, 1990
A;Title: Isolation and characterization of a human cDNA clone encoding a novel DNA topoi
A;Reference number: S10710; MUID:90306333; PMID:2163884
A;Accession: S10710
A;Molecule type: mRNA
A;Residues: 1043-1276 <AUS>
A;Cross-references: GB:X53662; GB:S56813; NID:g38324; PIDN:CAA37706.1; PID:g38325
R;Austin, C.A.; Sng, J.H.; Patel, S.; Fisher, L.M.
Biochim. Biophys. Acta 1172, 283-291, 1993
A;Title: Novel HeLa topoisomerase II is the II-beta isoform: complete coding sequenc
A;Reference number: S30190; MUID:93192319; PMID:8383537
A;Accession: S33970
A;Molecule type: mRNA
A;Residues: 1-23,29-1610,'A',1612-1626 <AU2>
A;Cross-references: EMBL:Z15111
R;Austin, C.A.; Sng, J.H.; Patel, S.; Fisher, L.M.
submitted to the EMBL Data Library, September 1992
A;Reference number: S30191
A;Accession: S30191
A;Molecule type: mRNA
A;Residues: 596-1430,'S',2432-1610,'A',1612-1626 <AU1>
A;Cross-references: EMBL:Z15115; NID:g288564; PIDN:CAA78821.1; PID:g288565
R;Davies, S.L.; Jenkins, J.R.; Hickson, I.D.
Nucleic Acids Res. 21, 3719-3723, 1993
A;Title: Human cells express two differentially spliced forms of topoisomerase II-be
A;Reference number: S41641; MUID:93376494; PMID:8396237
A;Accession: S41641
A;Molecule type: DNA
A;Residues: 24-80 <DAV>
A;Cross-references: EMBL:X71911; NID:g396540; PIDN:CAA50726.1; PID:g3980296
A;Note: this sequence represents a long minor splice form, designated beta-2
C;Genetics:
A;Gene: GDB:TOP2B
A;Cross-references: GDB:131575; OMIM:126431
A;Map position: 3p24-3p24
C;Superfamily: eukaryotic type II DNA topoisomerase; phage T4 DNA topoisomerase (ATP
C;Keywords: alternative splicing; ATP; dimer; isomerase; nucleus
F;1-1625/Product: DNA topoisomerase II beta-2 #status predicted <MINR>
F;1-23,29-1626/Product: DNA topoisomerase II beta-1 #status predicted <MAJR>
F;709-939/Domain: phage T4 DNA topoisomerase (ATP-hydrolyzing) medium chain homology

  Query Match             73.9%;  Score 34;  DB 2;  Length 1626;
  Best Local Similarity   66.7%;  Pred. No. 2.1e-02;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 YDDQESVKS 9
              :|| |:|:|
Db        543 YDDAESLKT 551


RESULT 14
C81674
hemolysin, probable TC0706 [imported] - Chlamydia muridarum (strain Nigg)
C;Species: Chlamydia muridarum, Chlamydia trachomatis MoPn
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 11-May-2000
C;Accession: C81674
R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hick
, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzb
Nucleic Acids Res. 28, 1397-1406, 2000
A;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR3
A;Reference number: A81500; MUID:20150255; PMID:10684935
A;Accession: C81674
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-374 <TET>
A;Cross-references: GB:AE002339; GB:AE002160; NID:g7190728; PIDN:AAF39520.1; PID:g71
A;Experimental source: strain Nigg (MoPn)
C;Genetics:
A;Gene: TC0706

  Query Match             71.7%;  Score 33;  DB 2;  Length 374;
  Best Local Similarity   85.7%;  Pred. No. 67;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 YDDQESV 7
              ||||| |
Db        277 YDDQEDV 283
```

```
RESULT 15
F64835
paraquat-inducible protein pqiB - Escherichia coli (strain K-12)
C;Species: Escherichia coli
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 01-Mar-2002
C;Accession: F64835; B57146; S52912
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: F64835
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-546 <BLAT>
A;Cross-references: GB:AE000197; GB:U00096; NID:g1787180; PIDN:AAC74037.1; PID:g1787184;
A;Experimental source: strain K-12, substrain MG1655
R;Koh, Y.S.; Roe, J.H.
J. Bacteriol. 177, 2673-2678, 1995
A;Title: Isolation of a novel paraquat-inducible (pqi) gene regulated by the soxRS locus
A;Reference number: A57146; MUID:95270582; PMID:7751279
A;Accession: B57146
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-91 <KOH>
A;Cross-references: EMBL:X81561
A;Experimental source: strain K-12, substrain W3110
C;Comment: Residues 1-282 appear to be homologous to PIR:B64945 and other members of tha
C;Genetics:
A;Gene: pqiB; pqi5B
C;Superfamily: pqiB protein
C;Keywords: transmembrane protein
F;19-35/Domain: transmembrane #status predicted <TMM>

  Query Match             71.7%;  Score 33;  DB 1;  Length 546;
  Best Local Similarity   62.5%;  Pred. No. 1e-02;
  Matches    5;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YDDQRSVK 8
              |||:|::
Db        271 YDDQKSIQ 278


Search completed: September 24, 2003, 13:49:20
Job time : 12.9213 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 24, 2003, 13:34:08 ; Search time 15.7753 Seconds
(without alignments)
79.250 Million cell updates/sec

Title: US-09-744-282-5
Perfect score: 65
Sequence: 1 PTPGRSFTLASSE 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	65	100.0	383	2	A49562	cartilage glycopro
2	62	95.4	405	2	S61551	breast-regressing
3	61	93.8	383	2	S51327	heparin-binding gl
4	45	69.2	525	2	T44445	chitinase (EC 3.2.
5	41	63.1	203	2	C25511	Cc protein - fruit
6	40	61.5	415	2	C87328	conserved hypothet
7	40	61.5	554	2	A56596	chitinase (EC 3.2.
8	40	61.5	1635	2	T14075	chitinase (EC 3.2.
9	39	60.0	113	2	I51025	transcription fact
10	39	60.0	483	2	A53918	chitinase (EC 3.2.
11	39	60.0	1199	2	A40670	nuclear envelope p
12	38	58.5	567	2	T49942	hypothetical prote
13	37	56.9	177	1	A69473	conserved hypothet
14	37	56.9	205	2	G69379	conserved hypothet
15	37	56.9	610	2	T39485	probable cystathio
16	37	56.9	641	2	T05497	hypothetical prote
17	37	56.9	684	2	B69308	conserved hypothet
18	37	56.9	1074	2	F72217	conserved hypothet
19	36	55.4	462	2	H70593	hypothetical prote
20	36	55.4	462	2	A75435	protoporphyrinogen
21	36	55.4	617	2	T15408	hypothetical prote
22	36	55.4	654	2	I38605	oviductal glycopro
23	36	55.4	676	2	AE2417	hypothetical prote
24	36	55.4	688	2	E86197	protein him-14 [im
25	36	55.4	765	2	E84553	hypothetical prote
26	35	53.8	175	2	H97174	uncharacterized pr
27	35	53.8	179	2	AG1994	hypothetical prote
28	35	53.8	364	2	C71295	probable octapreny
29	35	53.8	386	2	AG0081	conserved hypothet
30	35	53.8	387	2	S56411	hypothetical 45K p
31	35	53.8	387	2	B91274	probable synthetas
32	35	53.8	387	2	B86115	probable synthetas
33	35	53.8	387	2	AD0890	conserved hypothet
34	35	53.8	387	2	AH1050	conserved hypothet
35	35	53.8	402	2	AD3370	NAD(FAD)-utilizing
36	35	53.8	434	2	H83725	hypothetical prote
37	35	53.8	527	2	I15606	hypothetical prote
38	35	53.8	537	2	S57197	oviduct-specific g
39	35	53.8	539	2	I46470	estrogen dependent
40	35	53.8	550	2	T01770	hypothetical prote
41	35	53.8	591	2	S73790	hypothetical prote
42	35	53.8	849	2	H86430	T5I8.2 protein - A
43	35	53.8	1148	2	S51855	hypothetical prote
44	35	53.8	3716	2	E70969	probable PPE prote
45	34	52.3	208	2	D83895	hypothetical prote

ALIGNMENTS

RESULT 1
A49562
cartilage glycoprotein gp39 precursor - human
N;Alternate names: 39K synovial protein
C;Species: Homo sapiens (man)
C;Date: 23-Mar-1995 #sequence_revision 23-Mar-1995 #text_change 22-Jun-1999
C;Accession: A49562; S10677; A33162
R;Hakala, B.E.; White, C.; Recklies, A.D.
J. Biol. Chem. 268, 25803-25810, 1993
A;Title: Human cartilage gp-39, a major secretory product of articular chondrocytes
A;Reference number: A49562; MUID:94064658; PMID:8245017
A;Accession: A49562
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-383 <HAK>
A;Cross-references: GB:M80927; NID:g348911; PIDN:AAA16074.1; PID:g348912
R;Nyirkos, P.; Golds, E.E.
Biochem. J. 269, 265-268, 1990
A;Title: Human synovial cells secrete a 39 kDa protein similar to a bovine mammary
A;Reference number: S10677; MUID:90328983; PMID:2375755
A;Accession: S10677
A;Molecule type: protein
A;Residues: 22-40,'X',42-45 <NY2>
C;Superfamily: Streptomyces chitinase chi40
C;Keywords: cartilage; extracellular protein; glycoprotein
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-383/Product: cartilage glycoprotein gp39 #status predicted <MAT>

Query Match 100.0%; Score 65; DB 2; Length 383;
Best Local Similarity 100.0%; Pred. No. 0.00011;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PTPGRSFTLASSE 13
|||||||||||||
Db 259 PTPGRSFTLASSE 271

RESULT 2
S61551
breast-regressing protein brp39 precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 23-Jun-2000
C;Accession: S61551; S61550; I48271
R;Morrison, B.W.; Leder, P.
Oncogene 9, 3417-3426, 1994
A;Title: neu and ras initiate murine mammary tumors that share genetic markers gene
A;Reference number: I48271; MUID:95060797; PMID:7970700
A;Accession: S61551
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-405 <MOR1>
A;Cross-references: EMBL:X93035; NID:g1085065; PIDN:CAA63603.1; PID:g1085066

```
R;Morrison, B.W.
submitted to the EMBL Data Library, November 1995
A;Reference number: S61550
A;Accession: S61550
A;Molecule type: mRNA
A;Residues: 1-245,'I',247-330,'H',332-350,'MVWALDLDDFQGTCQPKEFFPLTNAIKDALA' <MOR2>
A;Cross-references: EMBL:X93035; NID:g1085065; PIDN:CAA63603.1; PID:g1085066
A;Note: the differences at the carboxyl end are due to a frameshift error
C;Genetics:
A;Gene: brp39
C;Superfamily: Streptomyces chitinase chi40
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-405/Product: breast-regressing protein brp39 #status predicted <MAT>

  Query Match             95.4%;  Score 62;  DB 2;  Length 405;
  Best Local Similarity   92.3%;  Pred. No. 0.00045;
  Matches   12;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PTFGRSFTLASSE 13
              |||||:|:|||||
Db        260 PTFGKSFTLASSE 272


RESULT 3
S51327
heparin-binding glycoprotein 38K - pig
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 22-Jun-1999
C;Accession: S51327
R;Shackelton, L.M.; Mann, D.M.; Millis, A.J.T.
submitted to the EMBL Data Library, January 1995
A;Description: Identification of a 38kDa heparin-binding glycoprotein (gp38k) in differe
A;Reference number: S51327
A;Accession: S51327
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-383 <SHA>
A;Cross-references: EMBL:Z47803; NID:g634097; PIDN:CAA87764.1; PID:g634098
C;Superfamily: Streptomyces chitinase chi40

  Query Match             93.8%;  Score 61;  DB 2;  Length 383;
  Best Local Similarity   92.3%;  Pred. No. 0.00066;
  Matches   12;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PTFGRSFTLASSE 13
              |||||||||:|||
Db        259 PTFGRSFTLASSK 271


RESULT 4
T44445
chitinase (EC 3.2.1.14) [imported] - African malaria mosquito
C;Species: Anopheles gambiae (African malaria mosquito)
C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jan-2000
C;Accession: T44445
R;Shen, Z.; Jacobs-Lorena, M.
submitted to the EMBL Data Library, June 1997
A;Reference number: Z22771
A;Accession: T44445
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-525 <SHE>
A;Cross-references: EMBL:AF008575; PIDN:AAB87764.1
A;Experimental source: adult; gut
C;Genetics:
A;Gene: chi-1
C;Keywords: glycosidase; hydrolase

  Query Match             69.2%;  Score 45;  DB 2;  Length 525;
  Best Local Similarity   69.7%;  Pred. No. 1.2;
  Matches    8;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 PTFGRSFTLASS 12
              ! :|:|:|||||:
Db        277 PLYGRNFTLASA 288


RESULT 5
C25511
Cc protein - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 21-Jul-2000
C;Accession: C25511
R;Eveleth Jr., D.D.; Marsh, J.L.
Nucleic Acids Res. 14, 6169-6183, 1986
A;Title: Sequence and expression of the Cc gene, a member of the dopa decarboxylase
A;Reference number: A93634; MUID:86312867; PMID:3092183
A;Accession: C25511
A;Molecule type: DNA
A;Residues: 1-203 <EVE>
A;Cross-references: GB:X04227; NID:g7635; PIDN:CAA27807.1; PID:g295740
C;Genetics:
A;Gene: FlyBase:l(2)37Cc
A;Cross-references: FlyBase:FBgn0002031
A;Introns: 67/3

  Query Match             63.1%;  Score 41;  DB 2;  Length 203;
  Best Local Similarity   88.9%;  Pred. No. 2.6;
  Matches    8;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 TFGRSFTLA 10
              |||| ||||
Db         97 TFGREFTLA 105


RESULT 6
C87328
conserved hypothetical protein CC0638 [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C;Accession: C87328
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg
 K.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; K
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser,
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: C87328
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-415 <STC>
A;Cross-references: GB:AE005673; NID:g13421851; PIDN:AAK22623.1; GSPDB:GN00148
C;Genetics:
A;Gene: CC0638

  Query Match             61.5%;  Score 40;  DB 2;  Length 415;
  Best Local Similarity   70.0%;  Pred. No. 8.7;
  Matches    7;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 PTFGRSFTLA 10
              ||||| | :|
Db         51 PTFGRKFLMA 60


RESULT 7
A56596
chitinase (EC 3.2.1.14) - tobacco hornworm
C;Species: Manduca sexta (tobacco hornworm)
C;Date: 11-Aug-1995 #sequence_revision 11-Aug-1995 #text_change 29-Jan-1999
C;Accession: A56596
R;Kramer, K.J.; Corpuz, L.; Choi, H.K.; Muthukrishnan, S.
Insect Biochem. Mol. Biol. 23, 691-701, 1993
A;Title: Sequence of a cDNA and expression of the gene encoding epidermal and gut ch
A;Reference number: A56596; MUID:93357793; PMID:8353525
A;Accession: A56596
A;Status: preliminary
```

A;Molecule type: mRNA
A;Residues: 1-554 <KRA>
A;Cross-references: GB:U02270; GB:S64757; NID:g406048; PID:g406049
A;Experimental source: larvae
A;Note: sequence extracted from NCBI backbone (NCBIN:136417, NCBIP:136418)
C;Keywords: glycosidase; hydrolase; polysaccharide degradation

  Query Match 61.5%; Score 40; DB 2; Length 554;
  Best Local Similarity 63.6%; Pred. No. 12;
  Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PTFGRSFTLAS 11
Db 268 PFYGRSFTLSA 278

RESULT 8
T14075
chitinase (EC 3.2.1.14) - yellow fever mosquito
C;Species: Aedes aegypti (yellow fever mosquito)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C;Accession: T14075
R;de la Vega, H.; Specht, C.A.; Liu, Y.; Robbins, P.W.
Insect Mol. Biol. 7, 233-239, 1997
A;Title: Chitinases are a multi-gene family in Aedes, Anopheles, and Drosophila.
A;Reference number: Z17872
A;Accession: T14075
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1635 <DEL>
A;Cross-references: EMBL:AF026492; NID:g2564720; PID:g2564721; PIDN:AAB81850.1
C;Genetics:
A;Gene: CHI2
A;Introns: 462/3; 524/3; 618/1; 951/3; 1151/2
C;Keywords: glycosidase; hydrolase; polysaccharide degradation

  Query Match 61.5%; Score 40; DB 2; Length 1635;
  Best Local Similarity 53.8%; Pred. No. 38;
  Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 PTFGRSFTLASSE 13
Db 484 PLYGQSFTLADQK 496

RESULT 9
I51025
transcription factor - eastern newt (fragment)
C;Species: Notophthalmus viridescens, Triturus viridescens (eastern newt)
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 24-Sep-1999
C;Accession: I51025
R;Simon, H.G.; Nelson, C.; Goff, D.; Laufer, E.; Morgan, B.A.; Tabin, C.
Dev. Dyn. 202, 1-12, 1995
A;Title: Differential expression of myogenic regulatory genes and Msx-1 during dedifferen
A;Reference number: I51024; MUID:95218226; PMID:7703517
A;Accession: I51025
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-113 <SIM>
A;Cross-references: EMBL:X82837; NID:g577786; PIDN:CAA58044.1; PID:g577787
C;Genetics:
A;Gene: Msx-1
C;Superfamily: unassigned homeobox proteins; homeobox homology
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
F;1-45/Domain: homeobox homology (fragment) <HOX>

  Query Match 60.0%; Score 39; DB 2; Length 113;
  Best Local Similarity 72.7%; Pred. No. 3.3;
  Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 PTFGRSFTLAS 11
Db 64 PTFGLSFPLGS 74

RESULT 10
A53918
chitinase (EC 3.2.1.14) precursor - braconid wasp (Chelonus sp.)
C;Species: Chelonus sp.
C;Date: 28-Jul-1995 #sequence_revision 28-Jul-1995 #text_change 21-Jul-2000
C;Accession: A53918
R;Krishnan, A.; Nair, P.N.; Jones, D.
J. Biol. Chem. 269, 20971-20976, 1994
A;Title: Isolation, cloning, and characterization of new chitinase stored in active
A;Reference number: A53918; MUID:94342256; PMID:8063715
A;Accession: A53918
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-483 <KRI>
A;Cross-references: GB:U10422; NID:g533504; PIDN:AAA61639.1; PID:g533505
C;Keywords: glycosidase; hydrolase; polysaccharide degradation

  Query Match 60.0%; Score 39; DB 2; Length 483;
  Best Local Similarity 54.5%; Pred. No. 16;
  Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PTFGRSFTLAS 11
Db 263 PAYGKSFTLSN 273

RESULT 11
A40670
nuclear envelope protein POM 121 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 05-Nov-1999
C;Accession: A40670
R;Hallberg, E.; Wozniak, R.W.; Blobel, G.
J. Cell Biol. 122, 513-521, 1993
A;Title: An integral membrane protein of the pore membrane domain of the nuclear en
A;Reference number: A40670; MUID:93328754; PMID:8335683
A;Accession: A40670
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-2199 <HAL>
A;Cross-references: GB:Z21513; NID:g396746; PIDN:CAA79225.1; PID:g396747
F;803-807,845-849,956-960,1010-1014,1047-1051,1076-1080/Region: pentapeptide motif

  Query Match 60.0%; Score 39; DB 2; Length 2199;
  Best Local Similarity 58.3%; Pred. No. 43;
  Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 PTFGRSFTLASS 12
Db 951 PSFGSSFTFGNS 962

RESULT 12
T49942
hypothetical protein F17I14.250 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
C;Accession: T49942
R;Bevan, M.; Hilbert, H.; Braun, M.; Holzer, E.; Brandt, A.; Duesterhoett, A.; Banc
submitted to the Protein Sequence Database, April 2000
A;Reference number: Z24490
A;Accession: T49942
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-567 <BEV>
A;Cross-references: EMBL:AL353994; GSPDB:GN00063; ATSP:F17I14.250
A;Experimental source: cultivar Columbia; BAC clone F17I14
C;Genetics:
A;Gene: ATSP:F17I14.250
A;Map position: 5
A;Introns: 169/3; 193/3; 361/3; 451/3

```
  Query Match           58.5%;  Score 38;  DB 2;  Length 567;
  Best Local Similarity  77.8%;  Pred. No. 30;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 PTFGRSFTL 9
              |||||| |:
Db        506 PTFGRSSTM 514


RESULT 13
A69473
conserved hypothetical protein AF1786 - Archaeoglobus fulgidus
C;Species: Archaeoglobus fulgidus
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000
C;Accession: A69473
R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
 Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A;Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.
Smith, H.O.; Woese, C.R.; Venter, J.C.
A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
A;Reference number: A69250; MUID:98049343; PMID:9389475
A;Accession: A69473
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-177 <KLE>
A;Cross-references: GB:AE000979; GB:AE000782; NID:g2689302; PIDN:AAB89461.1; PID:g264875
C;Superfamily: Bacillus subtilis hypothetical protein ywrF

  Query Match           56.9%;  Score 37;  DB 1;  Length 177;
  Best Local Similarity  58.3%;  Pred. No. 13;
  Matches    7;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 TFGRSFTLASSE 13
              :|||:|| |:|
Db        162 SFGKNFTTNSSE 173


RESULT 14
G69379
conserved hypothetical protein AF1039 - Archaeoglobus fulgidus
C;Species: Archaeoglobus fulgidus
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 18-Aug-2000
C;Accession: G69379
R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
 Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A;Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.
Smith, H.O.; Woese, C.R.; Venter, J.C.
A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
A;Reference number: A69250; MUID:98049343; PMID:9389475
A;Accession: G69379
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-205 <KLE>
A;Cross-references: GB:AE001031; GB:AE000782; NID:g2689354; PIDN:AAB90201.1; PID:g264959
C;Superfamily: Halobacterium salinarum taxis protein cheJ

  Query Match           56.9%;  Score 37;  DB 2;  Length 205;
  Best Local Similarity  60.0%;  Pred. No. 15;
  Matches    6;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 PTFGRSFTLA 10
              |:||| | :|
Db        142 PSFGRDFAIA 151


RESULT 15
T39485
probable cystathionine gamma-synthase - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C;Accession: T39485
R;Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Lucas, M.; Gaillardin, C.
submitted to the EMBL Data Library, August 1997
A;Reference number: Z21858
A;Accession: T39485
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-610 <LYN>
A;Cross-references: EMBL:AL031349; PIDN:CAA20484.1; GSPDB:GN00067; SPDB:SPBC15D4.09c
A;Experimental source: strain 972h-; cosmid c15D4
C;Genetics:
A;Gene: SPDB:SPBC15D4.09c
A;Map position: 2

  Query Match           56.9%;  Score 37;  DB 2;  Length 610;
  Best Local Similarity  63.6%;  Pred. No. 51;
  Matches    7;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 PTFGRSFTLAS 11
              |: | ::|||
Db        555 PSLGINFTLAS 565


Search completed: September 24, 2003, 13:49:23
Job time : 16.7753 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 24, 2003, 13:26:37 ; Search time 8.17978 Seconds
(without alignments)
74.739 Million cell updates/sec

Title: US-09-744-282-5
Perfect score: 65
Sequence: 1 PTPGKSFILASSE 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026795 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	65	100.0	383	1	C3L1_HUMAN	P36222 homo sapien
2	62	95.4	381	1	C3L1_MOUSE	Q61362 mus musculu
3	54	83.1	390	1	C3L2_HUMAN	Q15782 homo sapien
4	41	63.1	203	1	L2CC_DROME	P24156 drosophila
5	40	61.5	554	1	CHIT_MANSE	P36362 manduca sex
6	39	60.0	1199	1	P121_RAT	P52591 rattus norv
7	39	60.0	1229	1	P121_HUMAN	Q9y2n3 homo sapien
8	37	56.9	177	1	YH86_ARCFU	O28488 archaeoglob
9	37	56.9	610	1	MET7_SCHPO	O74314 schizosacch
10	37	56.9	816	1	CDB1_HUMAN	Q9y513 homo sapien
11	36	55.4	527	1	OGP_PIG	Q28990 sus scrofa
12	36	55.4	617	1	CHIT_CAEEL	Q11174 caenorhabdi
13	36	55.4	623	1	OGP_PAPAN	P36718 papio anubi
14	36	55.4	678	1	OGP_HUMAN	Q12889 homo sapien
15	36	55.4	715	1	ZDH5_HUMAN	Q9c0b5 homo sapien
16	36	55.4	715	1	ZDH5_MOUSE	Q8vdz4 mus musculu
17	36	55.4	842	1	HI14_CAEEL	Q23405 caenorhabdi
18	36	55.4	1997	1	OTOF_HUMAN	Q9hc10 homo sapien
19	36	55.4	1997	1	OTOF_MOUSE	Q9esf1 mus musculu
20	35	53.8	387	1	YJFC_ECOLI	P33222 escherichia
21	35	53.8	537	1	OGP_BOVIN	Q28042 bos taurus
22	35	53.8	539	1	OGP_SHEEP	Q28542 ovis aries
23	35	53.8	591	1	YD72_MYCPN	P75409 mycoplasma
24	35	53.8	721	1	OGP_MOUSE	Q62010 mus musculu
25	35	53.8	747	1	ORPB_HUMAN	Q9bxb4 homo sapien
26	34	52.3	295	1	YDH2_XANAU	P22644 xanthobacte
27	34	52.3	435	1	HISX_BUCAI	P57201 buchnera ap
28	34	52.3	453	1	ARX_BRARE	O42115 brachydanio
29	34	52.3	477	1	INGR_MOUSE	P15261 mus musculu
30	34	52.3	695	1	CMC1_DROME	Q9va73 drosophila
31	34	52.3	798	1	CDBE_HUMAN	Q9y5e9 homo sapien
32	34	52.3	1012	1	POLG_IBDVO	P27276 avian infec
33	34	52.3	1159	1	DP3A_VIBCH	P52022 vibrio chol
34	34	52.3	1190	1	ZO2_HUMAN	Q9udy2 homo sapien
35	34	52.3	1306	1	MSB2_YEAST	P32334 saccharomyc
36	34	52.3	1385	1	FAT1_SCHPO	Q13735 schizosacch
37	33	50.8	120	1	RET3_CHICK	P40220 gallus gall
38	33	50.8	195	1	TNF7_MOUSE	O55237 mus musculu
39	33	50.8	215	1	CIB3_HUMAN	Q9ny72 homo sapien
40	33	50.8	218	1	EST2_PSEFL	Q53547 pseudomonas
41	33	50.8	224	1	SSAK_SALTY	P74853 salmonella
42	33	50.8	268	1	TRPA_HAEIN	P43759 haemophilus
43	33	50.8	304	1	NODI_RHIS3	P72335 rhizobium s
44	33	50.8	319	1	AA3R_MOUSE	Q61618 mus musculu
45	33	50.8	351	1	VORB_METTH	O26800 methanobact

ALIGNMENTS

RESULT 1
C3L1_HUMAN
ID C3L1_HUMAN STANDARD; PRT; 383 AA.
AC P36222; P30923;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Chitinase-3 like protein 1 precursor (Cartilage glycoprotein-39) (GP-
DE 39) (39 kDa synovial protein) (YKL-40).
GN CHI3L1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Cartilage;
RX MEDLINE=94064658; PubMed=8245017;
RA Hakala B.E., White C., Recklies A.D.;
RT "Human cartilage gp-39, a major secretory product of articular
RT chondrocytes and synovial cells, is a mammalian member of a chitinase
RT protein family.";
RL J. Biol. Chem. 268:25803-25810(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=97386591; PubMed=9244440;
RA Rehli M., Krause S.W., Andreesen R.;
RT "Molecular characterization of the gene for human cartilage gp-39
RT (CHI3L1), a member of the chitinase protein family and marker for
RT late stages of macrophage differentiation.";
RL Genomics 43:221-225(1997).
RN [3]
RP SEQUENCE OF 22-45.
RX MEDLINE=90328983; PubMed=2375755;
RA Nyirkos P., Golds E.E.;
RT "Human synovial cells secrete a 39 kDa protein similar to a bovine
RT mammary protein expressed during the non-lactating period.";
RL Biochem. J. 269:265-268(1990).
CC -!- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN THE CAPACITY OF CELLS TO
CC RESPOND TO AND COPE WITH CHANGES IN THEIR ENVIRONMENT.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- TISSUE SPECIFICITY: PRESENT IN ARTICULAR CHONDROCYTES, SYNOVIAL
CC CELLS AS WELL AS IN LIVER. UNDETECTABLE IN MUSCLE TISSUES, LUNG,
CC PANCREAS, MONONUCLEAR CELLS, OR FIBROBLASTS.
CC -!- PTM: GLYCOSYLATED.
CC -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.
CC ---

CC ---

```
DR   EMBL; M80927; AAA16074.1; -.
DR   EMBL; Y08374; CAA69661.1; -.
DR   EMBL; Y08375; CAA69661.1; JOINED.
DR   EMBL; Y08376; CAA69661.1; JOINED.
DR   EMBL; Y08377; CAA69661.1; JOINED.
DR   EMBL; Y08378; CAA69661.1; JOINED.
DR   PIR; A49562; A49562.
DR   PDB; 1LA7; 10-APR-02.
DR   Genew; HGNC:1932; CHI3L1.
DR   MIM; 601525; -.
DR   GO; GO:0005578; C:extracellular matrix; TAS.
DR   GO; GO:0005615; C:extracellular space; TAS.
DR   GO; GO:0005207; F:extracellular matrix glycoprotein; TAS.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   InterPro; IPR001579; Glyco_hydro_18/2.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   ProDom; PD000471; Glyco_hydro_18; 1.
DR   SMART; SM00636; Glyco_18; 1.
DR   PROSITE; PS01095; CHITINASE_18; FALSE_NEG.
KW   Glycoprotein; Signal; 3D-structure.
FT   SIGNAL        1     21
FT   CHAIN        22    383       CHITINASE-3 LIKE PROTEIN 1.
FT   CARBOHYD     60     60       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   383 AA;  42613 MW;  76ADD8298EE8C2D1 CRC64;

  Query Match             100.0%;  Score 65;  DB 1;  Length 383;
  Best Local Similarity   100.0%;  Pred. No. 4.2e-05;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PTFGRSFTLASSE 13
              |||||||||||||
Db        259 PTFGRSFTLASSE 271


RESULT 2
C3L1_MOUSE
ID   C3L1_MOUSE     STANDARD;      PRT:   381 AA.
AC   Q61362;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Chitinase-3 like protein 1 precursor (Cartilage glycoprotein-39)
DE   (GP-39) (BRP39 protein).
GN   CHI3L1 OR BRP39.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muricae; Muridae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FVB/N; TISSUE=Breast;
RX   MEDLINE=95060797; PubMed=7970700;
RA   Morrison B.W., Leder P.;
RT   "neu and ras initiate murine mammary tumors that share genetic markers
RT   generally absent in c-myc and int-2-initiated tumors.";
RL   Oncogene 9:3417-3426(1994).
CC   -!- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN THE CAPACITY OF CELLS TO
CC       RESPOND TO AND COPE WITH CHANGES IN THEIR ENVIRONMENT (BY
CC       SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Extracellular (By similarity).
CC   -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X93035; CAA63603.1; -.
DR   PIR; S61551; S61551.
DR   MGD; MGI:1340899; Chi3l1.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   InterPro; IPR001579; Glyco_hydro_18/2.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   ProDom; PD000471; Glyco_hydro_18; 1.
DR   SMART; SM00636; Glyco_18; 1.
DR   PROSITE; PS01095; CHITINASE_18; FALSE_NEG.
KW   Glycoprotein; Signal.
FT   SIGNAL        1     21       POTENTIAL.
FT   CHAIN        22    381       CHITINASE-3 LIKE PROTEIN 1.
FT   CARBOHYD     60     60       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   381 AA;  43061 MW;  EF6581E8184F0450 CRC64;

  Query Match             95.4%;  Score 62;  DB 1;  Length 381;
  Best Local Similarity   92.3%;  Pred. No. 0.00016;
  Matches   12;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PTFGRSFTLASSE 13
              ||||:|||||||
Db        260 PTFGKSFTLASSE 272


RESULT 3
C3L2_HUMAN
ID   C3L2_HUMAN     STANDARD;      PRT:   390 AA.
AC   Q15782; Q15749; Q15783;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Chitinase 3-like protein 2 precursor (YKL-39) (Chondrocyte protein
DE   39).
GN   CHI3L2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS LONG AND SHORT).
RA   Grossman A., Matsuyama T., Baker E., Waterhouse P., Sutherland G.R.,
RA   Mak T.W.;
RT   "Cloning of a novel lymphoid restricted human chitinase and
RT   localization to 1p13.3.";
RL   Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM SHORT).
RC   TISSUE=Articular cartilage;
RX   MEDLINE=96325055; PubMed=8702629;
RA   Hu B., Trinh K., Figueira W.F., Price P.A.;
RT   "Isolation and sequence of a novel human chondrocyte protein related
RT   to mammalian members of the chitinase protein family.";
RL   J. Biol. Chem. 271:19415-19420(1996).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=Short;
CC         IsoId=Q15782-1; Sequence=Displayed;
CC       Name=Long;
CC         IsoId=Q15782-2; Sequence=VSP_001066;
CC         Note=No experimental confirmation available;
CC   -!- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN CHONDROCYTES, FOLLOWED
CC       BY SYNOVIOCYTES, LUNG AND HEART. NOT DETECTED IN BRAIN, SPLEEN,
CC       PANCREAS, AND LIVER.
CC   -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U58515; AAB04534.1; -.
DR   EMBL; U58514; AAB04533.1; -.
DR   EMBL; U49835; AAC50597.1; ALT_INIT.
```

```
DR   Genew; HGNC:1933; CHI3L2.
DR   MIM; 601526; -.
DR   GO; GO:0005615; C:extracellular space; TAS.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   InterPro; IPR001579; Glyco_hydro_18/2.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   ProDom; PD000471; Glyco_hydro_18; 1.
DR   SMART; SM00636; Glyco_18; 1.
DR   PROSITE; PS01095; CHITINASE_18; FALSE_NEG.
KW   Hydrolase; Signal; Alternative splicing.
FT   SIGNAL        1     26       POTENTIAL.
FT   CHAIN        27    390       CHITINASE 3-LIKE PROTEIN 2.
FT   CARBOHYD     35     35       N-LINKED (GLCNAC...) (POTENTIAL).
FT   VARSPLIC      1     12       MGAITMDQKSLW -> SQQIWLGSPDEVWLYLVCRHTHIFI
FT                                AIWDSRTRFGSTDFSFHPSI (in isoform Long).
FT                                /FTId=VSP_001066.
SQ   SEQUENCE   390 AA;  43500 MW;  97B86A2F3AA35677 CRC64;

  Query Match             88.1%;  Score 54;  DB 1;  Length 390;
  Best Local Similarity   76.9%;  Pred. No. 0.0061;
  Matches   10;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 PTFGRSFTLASSE 13
              |||| |||||:||
Db        267 PTYGHSFTLASAE 279


RESULT 4
L2CC_DROME
ID   L2CC_DROME     STANDARD;      PRT;   203 AA.
AC   P24156;
DT   01-MAR-1992 (Rel. 21, Created)
DT   01-MAR-1992 (Rel. 21, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   L(2)37CC protein.
GN   L(2)37CC.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86312887; PubMed=3092153;
RA   Eveleth D.D. Jr., Marsh J.L.;
RT   "Sequence and expression of the Cc gene, a member of the dopa
RT   decarboxylase gene cluster of Drosophila: possible translational
RT   regulation.";
RL   Nucleic Acids Res. 14:6169-6183(1986).
CC   -!- FUNCTION: REQUIRED FOR LARVAL METABOLISM OR FOR THE PROGRESSION
CC       OF THE LARVA INTO A PUPA.
CC   -!- SIMILARITY: BELONGS TO THE PROHIBITIN FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X04228; CAA27810.1; -.
DR   EMBL; X04227; CAA27807.1; -.
DR   PIR; C25511; C25511.
DR   FlyBase; FBgn0002031; l(2)37Cc.
DR   InterPro; IPR001107; Band_7.
DR   InterPro; IPR000163; Prohibitin.
DR   Pfam; PF01145; Band_7; 1.
DR   PRINTS; PR00679; PROHIBITIN.
DR   SMART; SM00244; PHB; 1.
KW   Developmental protein.
SQ   SEQUENCE   203 AA;  22817 MW;  B15D085CC0862A11 CRC64;

  Query Match             63.1%;  Score 41;  DB 1;  Length 203;
  Best Local Similarity   88.9%;  Pred. No. 1;
  Matches    8;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 TFGRSFTLA 10
              |||| ||||
Db         97 TFGREFTLA 105


RESULT 5
CHIT_MANSE
ID   CHIT_MANSE     STANDARD;      PRT;   554 AA.
AC   P36362;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-FEB-1995 (Rel. 31, Last annotation update)
DE   Endochitinase precursor (EC 3.2.1.14).
OS   Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Sphingoidea;
OC   Sphingidae; Sphinginae; Manduca.
OX   NCBI_TaxID=7130;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93357793; PubMed=8353525;
RA   Kramer K.J., Corpuz L., Choi H.K., Muthukrishnan S.;
RT   "Sequence of a cDNA and expression of the gene encoding epidermal and
RT   gut chitinases of Manduca sexta.";
RL   Insect Biochem. Mol. Biol. 23:691-701(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97215580; PubMed=9061927;
RA   Choi H.K., Choi K.H., Kramer K.J., Muthukrishnan S.;
RT   "Isolation and characterization of a genomic clone for the gene of an
RT   insect molting enzyme, chitinase.";
RL   Insect Biochem. Mol. Biol. 27:37-47(1997).
CC   -!- FUNCTION: DIGEST CHITIN IN THE EXOSKELETON DURING THE MOLTING
CC       PROCESS.
CC   -!- CATALYTIC ACTIVITY: Hydrolysis of the 1,4-beta-linkages of N-
CC       acetyl-D-glucosamine polymers of chitin.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: EPIDERMIS AND GUT.
CC   -!- DEVELOPMENTAL STAGE: HIGH LEVELS SEEN IN THE EPIDERMIS ON DAY 0,
CC       BUT RAPIDLY DISAPPEARS AND IS UNDETECTED ON DAYS 1-4 OF FIFTH
CC       INSTAR. IT REAPPEARS ON DAY 5 AND PEAKS ON DAY 7 AFTER WHICH A
CC       RAPID DECLINE IS SEEN. IN THE GUT IS DETECTED ON DAY 6 WITH LOWER
CC       LEVELS SEEN ON DAYS 0, 7 AND 8.
CC   -!- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
CC       HYDROLASES).
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; U02270; AAC04924.1; -.
DR   EMBL; L49234; AAB53952.1; -.
DR   PIR; A56596; A56596.
DR   InterPro; IPR002557; Chitin_bind_PerA.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   InterPro; IPR001579; Glyco_hydro_18/2.
DR   Pfam; PF01607; CBM_14; 1.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   ProDom; PD000471; Glyco_hydro_18; 1.
DR   SMART; SM00494; ChtBD2; 1.
DR   SMART; SM00636; Glyco_18; 1.
DR   PROSITE; PS01095; CHITINASE_18; 1.
KW   Hydrolase; Glycosidase; Chitin degradation; Signal; Glycoprotein.
FT   SIGNAL        1     19       POTENTIAL.
FT   CHAIN        20    554       ENDOCHITINASE.
```

```
FT   DOMAIN      396    453       SER/THR-RICH.
FT   ACT_SITE    146    146       PROTON DONOR (BY SIMILARITY).
FT   CARBOHYD     85     85       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    303    303       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    407    407       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    545    545       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   554 AA;  62203 MW;  3989D756C96CD490 CRC64;

  Query Match             61.5%;  Score 40;  DB 1;  Length 554;
  Best Local Similarity   63.6%;  Pred. No. 5.1;
  Matches    7;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 PTFGRSFTLAS 11
              | :|||||||::
Db        268 PFYGRSFTLSA 278


RESULT 6
P121_RAT
ID   P121_RAT       STANDARD;      PRT;  1199 AA.
AC   P52591;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Nuclear envelope pore membrane protein POM 121 (Pore membrane protein
DE   of 121 kDa) (P145).
GN   POM121 OR NUP121.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC   STRAIN=Sprague-Dawley; TISSUE=Liver;
RX   MEDLINE=93328754; PubMed=8335683;
RA   Hallberg E., Wozniak R.W., Blobel G.;
RT   "An integral membrane protein of the pore membrane domain of the
RT   nuclear envelope contains a nucleoporin-like region.";
RL   J. Cell Biol. 122:513-521(1993).
CC   -!- FUNCTION: ESSENTIAL COMPONENT OF THE NUCLEAR PORE COMPLEX. THE
CC       REPEAT-CONTAINING DOMAIN MAY BE INVOLVED IN ANCHORING COMPONENTS
CC       OF THE PORE COMPLEX TO THE PORE MEMBRANE.
CC   -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. NUCLEAR PORE
CC       MEMBRANE.
CC   -!- DOMAIN: CONTAINS F-X-F-G REPEATS.
CC   -!- SIMILARITY: THE REPEAT REGION COMPOSED OF PENTAPEPTIDE REPEATS
CC       SEPARATED BY SER/THR-RICH DOMAINS IS SIMILAR TO THAT OF YEAST
CC       NUP1, NSP1 AND MAMMALIAN P62 AND NUP153.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z21513; CAA79725.1; -.
DR   EMBL; Z21514; CAA79726.1; -.
DR   PIR; A40670; A40670.
DR   Pfam; PF03093; Nucleoporin_FG; 8.
KW   Nuclear protein; Transport; Transmembrane; Repeat.
FT   DOMAIN        1     55       CISTERNAL SIDE (POTENTIAL).
FT   TRANSMEM     57     77       POTENTIAL.
FT   DOMAIN       76   1199       PORE SIDE (POTENTIAL).
FT   DOMAIN        4      8       POLY-ALA.
FT   DOMAIN       53     58       POLY-ALA.
FT   DOMAIN      438    441       POLY-SER.
FT   DOMAIN      497    500       POLY-PRO.
FT   DOMAIN      684    687       POLY-SER.
FT   DOMAIN     1023   1026       POLY-ALA.
FT   DOMAIN     1033   1038       POLY-SER.
FT   DOMAIN     1058   1063       POLY-GLY.
SQ   SEQUENCE   1199 AA;  120784 MW;  6DC4451B81D5B927 CRC64;

  Query Match             60.0%;  Score 39;  DB 1;  Length 1199;
  Best Local Similarity   58.3%;  Pred. No. 20;
  Matches    7;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 PTFGRSFTLASS 12
              |:|| ||| :|
Db        951 PSFGSSFTFGNS 962


RESULT 7
P121_HUMAN
ID   P121_HUMAN     STANDARD;      PRT;  1229 AA.
AC   Q9Y2N3; O75115; Q9Y4S7;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Nuclear envelope pore membrane protein POM 121 (Pore membrane protein
DE   of 121 kDa) (P145).
GN   POM121 OR NUP121 OR KIAA0618.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Cordes M., Bauer C., Holmes A.;
RL   Submitted (NOV-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE OF 243-1229 FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=98403880; PubMed=9734811;
RA   Ishikawa K.-I., Nagase T., Suyama M., Miyajima N., Tanaka A.,
RA   Kotani H., Nomura N., Ohara O.;
RT   "Prediction of the coding sequences of unidentified human genes. X.
RT   The complete sequences of 100 new cDNA clones from brain which can
RT   code for large proteins in vitro.";
RL   DNA Res. 5:169-176(1998).
RN   [3]
RP   SEQUENCE OF 1130-1229 FROM N.A.
RC   TISSUE=Uterus;
RA   Koehrer K., Beyer A., Mewes H.-W., Gassenhuber J., Wiemann S.;
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: ESSENTIAL COMPONENT OF THE NUCLEAR PORE COMPLEX. THE
CC       REPEAT-CONTAINING DOMAIN MAY BE INVOLVED IN ANCHORING COMPONENTS
CC       OF THE PORE COMPLEX TO THE PORE MEMBRANE (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. NUCLEAR PORE
CC       MEMBRANE (BY SIMILARITY).
CC   -!- DOMAIN: CONTAINS F-X-F-G REPEATS.
CC   -!- SIMILARITY: THE REPEAT REGION COMPOSED OF PENTAPEPTIDE REPEATS
CC       SEPARATED BY SER/THR-RICH DOMAINS IS SIMILAR TO THAT OF YEAST
CC       NUP1, NSP1 AND MAMMALIAN P62 AND NUP153.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AC006014; AAD28064.1; -.
DR   EMBL; AB014518; BAA31593.1; -.
DR   EMBL; AL080109; CAB45713.1; -.
DR   PIR; T12551; T12551.
DR   Genew; HGNC:19702; POM121.
KW   Nuclear protein; Transport; Transmembrane; Repeat.
FT   DOMAIN        1     40       CISTERNAL SIDE (POTENTIAL).
FT   TRANSMEM     41     61       POTENTIAL.
FT   DOMAIN       62   1229       PORE SIDE (POTENTIAL).
FT   DOMAIN        4     10       POLY-ALA.
FT   DOMAIN       51     56       POLY-ALA.
```

```
FT   DOMAIN       294     299       POLY-LYS.
FT   DOMAIN       443     444       POLY-SER.
FT   DOMAIN       499     502       POLY-PRO.
FT   DOMAIN       733     736       POLY-SER.
FT   DOMAIN       819     825       POLY-SER.
FT   DOMAIN       869     875       POLY-THR.
FT   DOMAIN      1062    1067       POLY-SER.
FT   CONFLICT     297     297       K -> E (IN REF. 2).
FT   CONFLICT     336     336       S -> N (IN REF. 2).
FT   CONFLICT     379     379       T -> A (IN REF. 2).
FT   CONFLICT     451     451       K -> R (IN REF. 2).
FT   CONFLICT     545     545       T -> I (IN REF. 2).
FT   CONFLICT     559     559       T -> P (IN REF. 2).
FT   CONFLICT     702     702       S -> P (IN REF. 2).
FT   CONFLICT     881     881       P -> H (IN REF. 2).
FT   CONFLICT     965     965       T -> A (IN REF. 2).
FT   CONFLICT     982     984       PAT -> A (IN REF. 2).
FT   CONFLICT     991     996       ASTIKI -> PSMIKV (IN REF. 2).
FT   CONFLICT    1000    1000       H -> Y (IN REF. 2).
FT   CONFLICT    1006    1008       QPT -> HPI (IN REF. 2).
FT   CONFLICT    1044    1044       MISSING (IN REF. 2).
FT   CONFLICT    1102    1102       A -> T (IN REF. 2).
FT   CONFLICT    1124    1124       T -> A (IN REF. 2).
FT   CONFLICT    1146    1146       S -> G (IN REF. 2).
FT   CONFLICT    1165    1165       L -> Q (IN REF. 1).
FT   CONFLICT    1195    1195       A -> G (IN REF. 1).
FT   CONFLICT    1199    1229       SAAPSFSIGAGSKTPGARQRLQARRQHTRKK -> NTFAHQ
FT                                  QEHSPRKGPNNLSKRKLLPAVRAQGPPRRGQASSFPIRKE
FT                                  (IN REF. 2).
FT   CONFLICT    1202    1202       P -> L (IN REF. 3).
SQ   SEQUENCE   1229 AA;  125087 MW;  515655D12858998B CRC64;

  Query Match             60.0%;  Score 39;  DB 1;  Length 1229;
  Best Local Similarity   58.3%;  Pred. No. 20;
  Matches    7;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 PTFGRSFTLASS 12
              |:|: ||:  :|
Db        966 PSFGSSFTFGNS 977


RESULT 8
YH86_ARCFU
ID   YH86_ARCFU     STANDARD;      PRT;   177 AA.
AC   O28488;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical protein AF1786.
GN   AF1786.
OS   Archaeoglobus fulgidus.
OC   Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC   Archaeoglobaceae; Archaeoglobus.
OX   NCBI_TaxID=2234;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX   MEDLINE=98049343; PubMed=9389475;
RA   Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA   Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA   Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,
RA   Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA   Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA   Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA   Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA   Cotton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,
RA   Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA   Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA   Venter J.C.;
RT   "The complete genome sequence of the hyperthermophilic, sulphate-
RT   reducing archaeon Archaeoglobus fulgidus.";
RL   Nature 390:364-370(1997).
CC   -!- COFACTOR: FMN (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE FLAVOREDOXIN FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AE000979; AAB89461.1; -.
DR   PIR; A69473; A69473.
DR   TIGR; AF1786; -.
DR   InterPro; IPR002563; Flavin_Reduct.
DR   Pfam; PF01613; Flavin_Reduct; 1.
KW   Hypothetical protein; Flavoprotein; FMN; Complete proteome.
SQ   SEQUENCE   177 AA;  19830 MW;  8E48DAF17A48C0E3 CRC64;

  Query Match             56.9%;  Score 37;  DB 1;  Length 177;
  Best Local Similarity   58.3%;  Pred. No. 5.3;
  Matches    7;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 TFGRSFTLASSE 13
              :|||::||  |||
Db        152 SFGKNFTINSSE 173


RESULT 9
MET7_SCHPO
ID   MET7_SCHPO     STANDARD;      PRT;   610 AA.
AC   O74314; Q9USA0;
DT   15-SEP-2003 (Rel. 42, Created)
DT   15-SEP-2003 (Rel. 42, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Probable cystathionine gamma-synthase (EC 4.2.99.9) (O-
DE   succinylhomoserine (Thiol)-lyase).
GN   SPBC15D4.09C.
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
OC   Schizosaccharomyces.
OX   NCBI_TaxID=4896;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972;
RX   MEDLINE=21848401; PubMed=11859360;
RA   Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA   Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA   Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA   Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA   Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA   Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA   James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA   Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA   Oliver K., O'Neil S., Pearson D., Quail M.A., Rabbinowitsch E.,
RA   Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA   Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA   Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA   Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA   Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA   Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,
RA   Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA   Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA   Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA   Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA   Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA   Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA   Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA   Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA   Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT   "The genome sequence of Schizosaccharomyces pombe.";
RL   Nature 415:871-880(2002).
RN   [2]
```

```
RP   SEQUENCE OF 1-213 FROM N.A., AND SUBCELLULAR LOCATION.
RC   STRAIN=968 h90;
RX   MEDLINE=20223868; PubMed=10759889;
RA   Ding D.-Q., Tomita Y., Yamamoto A., Chikashige Y., Haraguchi T.,
RA   Hiraoka Y.;
RT   "Large-scale screening of intracellular protein localization in living
RT   fission yeast cells by the use of a GFP-fusion genomic DNA library.";
RL   Genes Cells 5:169-190(2000).
CC   -!- CATALYTIC ACTIVITY: O-succinyl-L-homoserine + L-cysteine =
CC       cystathionine + succinate.
CC   -!- COFACTOR: Pyridoxal phosphate (By similarity).
CC   -!- PATHWAY: Methionine biosynthesis; second step.
CC   -!- SUBCELLULAR LOCATION: Nuclear, and cytoplasmic.
CC   -!- SIMILARITY: Belongs to the trans-sulfuration enzymes family.
CC       MET7 subfamily.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AL031349; CAA20484.1; -.
DR   EMBL; AB027923; BAA87227.1; -.
DR   GeneDB_SPombe; SPBC15D4.09c; -.
DR   InterPro; IPR000277; Cys_Met_Meta_PP.
DR   Pfam; PF01053; Cys_Met_Meta_PP; 1.
KW   Methionine biosynthesis; Lyase; Pyridoxal phosphate; Nuclear protein.
FT   BINDING     427    427       PYRIDOXAL PHOSPHATE (POTENTIAL).
SQ   SEQUENCE   610 AA;  68968 MW;  129F54765ADC71EF CRC64;

  Query Match             56.9%;  Score 37;  DB 1;  Length 610;
  Best Local Similarity   63.6%;  Pred. No. 22;
  Matches    7;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 PTFGRSFTLAS 11
              |: | :||||:
Db        555 PSLGTNPTLAS 565


RESULT 10
CDB1_HUMAN
ID   CDB1_HUMAN     STANDARD;      PRT;   818 AA.
AC   Q9Y5F3;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Protocadherin beta 1 precursor (PCDH-beta1).
GN   PCDHB1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99308636; PubMed=10380929;
RA   Wu Q., Maniatis T.;
RT   "A striking organization of a large family of human neural cadherin-
RT   like cell adhesion genes.";
RL   Cell 97:779-790(1999).
CC   -!- FUNCTION: POTENTIAL CALCIUM-DEPENDENT CELL-ADHESION PROTEIN.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC   -!- SIMILARITY: Contains 6 cadherin domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF152488; AAD43749.1; -.
DR   Genew; HGNC:8680; PCDHB1.
DR   MIM; 606327; -.
DR   MIM; 604967; -.
DR   InterPro; IPR002126; Cadherin.
DR   Pfam; PF00028; cadherin; 5.
DR   PRINTS; PR00205; CADHERIN.
DR   SMART; SM00112; CA; 6.
DR   PROSITE; PS00232; CADHERIN_1; 5.
DR   PROSITE; PS50268; CADHERIN_2; 6.
KW   Calcium-binding; Cell adhesion; Glycoprotein; Signal; Repeat;
KW   Transmembrane; Multigene family.
FT   SIGNAL        1     28       POTENTIAL.
FT   CHAIN        29    818       PROTOCADHERIN BETA 1.
FT   DOMAIN       29    691       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    692    712       POTENTIAL.
FT   DOMAIN      713    818       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       35    133       CADHERIN 1.
FT   DOMAIN      138    242       CADHERIN 2.
FT   DOMAIN      243    347       CADHERIN 3.
FT   DOMAIN      348    452       CADHERIN 4.
FT   DOMAIN      457    562       CADHERIN 5.
FT   DOMAIN      577    672       CADHERIN 6.
FT   CARBOHYD    169    169       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    209    209       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    257    257       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    419    419       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    568    568       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   818 AA;  90456 MW;  4EBF7EF4EDB7D604 CRC64;

  Query Match             56.9%;  Score 37;  DB 1;  Length 818;
  Best Local Similarity   54.5%;  Pred. No. 31;
  Matches    6;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 PTFGRSFTLAS 11
              |||' |::| :
Db        396 PTFGNSYSLVT 406


RESULT 11
OGP_PIG
ID   OGP_PIG        STANDARD;      PRT;   527 AA.
AC   Q28990;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Oviduct-specific glycoprotein precursor (Oviductal glycoprotein)
DE   (Oviductin) (Estrogen-dependent oviduct protein) (OSSP-E3).
GN   OVGP1 OR OGP.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Yorkshire X Duroc X Hampshire; TISSUE=Oviduct;
RX   MEDLINE=97107140; PubMed=8949888;
RA   Buhi W.C., Alvarez I.M., Choi I., Cleaver B.D., Simmen F.A.;
RT   "Molecular cloning and characterization of an estrogen-dependent
RT   porcine oviductal secretory glycoprotein.";
RL   Biol. Reprod. 55:1305-1314(1996).
CC   -!- FUNCTION: BINDS TO OOCYTE ZONA PELLUCIDA IN VIVO. MAY PLAY A ROLE
CC       IN THE FERTILIZATION PROCESS AND/OR EARLY EMBRYONIC DEVELOPMENT.
CC   -!- SUBCELLULAR LOCATION: Secretory granules.
CC   -!- TISSUE SPECIFICITY: OVIDUCT.
CC   -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
```

```
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ------------------------------------------------------------------------
DR   EMBL; U43490; AAA85445.1; -.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   InterPro; IPR001579; Glyco_hydro_18/2.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   ProDom; PD000471; Glyco_hydro_18; 1.
DR   SMART; SM00636; Glyco_18; 1.
DR   PROSITE; PS01095; CHITINASE_18; FALSE_NEG.
KW   Glycoprotein; Fertilization; Signal.
FT   SIGNAL       1     21       BY SIMILARITY.
FT   CHAIN       22    527       OVIDUCT-SPECIFIC GLYCOPROTEIN.
FT   CARBOHYD    62     62       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   402    402       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   441    441       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   527 AA;  58519 MW;  31B78F49CA2363A2 CRC64;

  Query Match           55.4%;  Score 36;  DB 1;  Length 527;
  Best Local Similarity   66.7%;  Pred. No. 29;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 PTFGRSFIL 9
              |||:|:|:| |
Db        261 PTYGRTFRL 269


RESULT 12
CHIT_CAEEL
ID   CHIT_CAEEL     STANDARD;      PRT;   617 AA.
AC   Q11174;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Probable endochitinase (EC 3.2.1.14).
GN   CHT-1 OR C04F6.3.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Nhan M.;
RL   Submitted (DEC-1995) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: Hydrolysis of the 1,4-beta-linkages of N-
CC       acetyl-D-glucosamine polymers of chitin.
CC   -!- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
CC       HYDROLASES).
CC   ------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ------------------------------------------------------------------------
DR   EMBL; U42835; AAA83586.1; -.
DR   PIR; T15408; T15408.
DR   WormPep; C04F6.3; CBG03923.
DR   InterPro; IPR002557; Chitin_bind_PerA.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   InterPro; IPR001579; Glyco_hydro_18/2.
DR   Pfam; PF01607; CBM_14; 2.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   ProDom; PD000471; Glyco_hydro_18; 1.
DR   SMART; SM00494; ChtBD2; 2.
DR   SMART; SM00636; Glyco_18; 1.
DR   PROSITE; PS01095; CHITINASE_18; 1.
KW   Hydrolase; Glycosidase; Chitin degradation.
FT   ACT_SITE    179    179       PROTON DONOR (BY SIMILARITY).
FT   DOMAIN      435    478       THR-RICH.
SQ   SEQUENCE   617 AA;  66857 MW;  DDA1D2AAAC0E54DA CRC64;

  Query Match           55.4%;  Score 36;  DB 1;  Length 617;
  Best Local Similarity   66.7%;  Pred. No. 35;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 PTFGRSFIL 9
              |||:|:|:| |
Db        300 PTYGRGWTL 308


RESULT 13
OGP_PAPAN
ID   OGP_PAPAN      STANDARD;      PRT;   623 AA.
AC   P36718;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Oviduct-specific glycoprotein precursor (Oviductal glycoprotein)
DE   (Oviductin) (Estrogen-dependent oviduct protein).
GN   OVGP1 OR OGP.
OS   Papio anubis (Olive baboon).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Papio.
OX   NCBI_TaxID=9555;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Oviduct;
RX   MEDLINE=98244335; PubMed=9584944;
RA   Verhage H.G., Fazleabas A.T., Mavrogianis P.A., O'Day-Bowman M.B.,
RA   Donnelly K.M., Arias E.B., Jaffe R.C.;
RT   "The baboon oviduct: characteristics of an oestradiol-dependent
RT   oviduct-specific glycoprotein.";
RL   Hum. Reprod. Update 3:541-552(1997).
RN   [2]
RP   PRELIMINARY PARTIAL SEQUENCE FROM N.A.
RC   TISSUE=Oviduct;
RX   MEDLINE=91367180; PubMed=1716345;
RA   Donnelly K.M., Fazleabas A.T., Verhage H.G., Mavrogianis P.A.,
RA   Jaffe R.C.;
RT   "Cloning of a recombinant complementary DNA to a baboon (Papio
RT   anubis) estradiol-dependent oviduct-specific glycoprotein.";
RL   Mol. Endocrinol. 5:356-364(1991).
CC   -!- FUNCTION: BINDS TO OOCYTE ZONA PELLUCIDA IN VIVO. MAY PLAY A ROLE
CC       IN THE FERTILIZATION PROCESS AND/OR EARLY EMBRYONIC DEVELOPMENT.
CC   -!- SUBCELLULAR LOCATION: Secretory granules.
CC   -!- TISSUE SPECIFICITY: OVIDUCT.
CC   -!- DEVELOPMENTAL STAGE: AT THE TIME OF OVULATION.
CC   -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.
CC   ------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ------------------------------------------------------------------------
DR   EMBL; M59903; AAB39765.1; -.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   InterPro; IPR001579; Glyco_hydro_18/2.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   ProDom; PD000471; Glyco_hydro_18; 1.
DR   SMART; SM00636; Glyco_18; 1.
DR   PROSITE; PS01095; CHITINASE_18; FALSE_NEG.
KW   Glycoprotein; Fertilization; Signal.
FT   SIGNAL       1     21       BY SIMILARITY.
FT   CHAIN       22    623       OVIDUCT-SPECIFIC GLYCOPROTEIN.
FT   CARBOHYD   402    402       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   441    441       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   623 AA;  69291 MW;  9E21CE481FFF1268 CRC64;
```

```
  Query Match             55.4%;  Score 36;  DB 1;  Length 623;
  Best Local Similarity   66.7%;  Pred. No. 36;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 PTFGRSFTL 9
              ||:|||:| |
Db        261 PTYGRTFRL 269


RESULT 14
OGP_HUMAN
ID   OGP_HUMAN      STANDARD;      PRT:   678 AA.
AC   Q12889; O15841;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Oviduct-specific glycoprotein precursor (Oviductal glycoprotein)
DE   (Oviductin) (Estrogen-dependent oviduct protein) (Mucin 9).
GN   OVGP1 OR OGP OR MUC9.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Oviduct;
RX   MEDLINE=95119256; PubMed=7819450;
RA   Arias E.B., Verhage H.G., Jaffe R.C.;
RT   "Complementary deoxyribonucleic acid cloning and molecular
RT   characterization of an estrogen-dependent human oviductal
RT   glycoprotein.";
RL   Biol. Reprod. 51:685-694(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Jaffe R.C.;
RL   Submitted (JUL-1996) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Coville G.;
RL   Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: BINDS TO OOCYTE ZONA PELLUCIDA IN VIVO. MAY PLAY A ROLE
CC       IN THE FERTILIZATION PROCESS AND/OR EARLY EMBRYONIC DEVELOPMENT.
CC   -!- SUBCELLULAR LOCATION: Secretory granules.
CC   -!- TISSUE SPECIFICITY: OVIDUCT.
CC   -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; U09550; AAA86946.1; -.
DR   EMBL; U58010; AAB04126.1; -.
DR   EMBL; U58001; AAB04126.1; JOINED.
DR   EMBL; U58002; AAB04126.1; JOINED.
DR   EMBL; U58003; AAB04126.1; JOINED.
DR   EMBL; U58004; AAB04126.1; JOINED.
DR   EMBL; U58005; AAB04126.1; JOINED.
DR   EMBL; U58006; AAB04126.1; JOINED.
DR   EMBL; U58007; AAB04126.1; JOINED.
DR   EMBL; U58008; AAB04126.1; JOINED.
DR   EMBL; U58009; AAB04126.1; JOINED.
DR   EMBL; AL390195; CAC36039.1; -.
DR   Genew; HGNC:8524; OVGP1.
DR   MIM; 603578; -.
DR   GO; GO:0007565; P:pregnancy; IAS.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   InterPro; IPR001579; Glyco_hydro_18/2.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   ProDom; PD000471; Glyco_hydro_18; 1.
DR   SMART; SM00636; Glyco_18; 1.
DR   PROSITE; PS01095; CHITINASE_18; FALSE_NEG.
KW   Glycoprotein; Fertilization; Signal.
FT   SIGNAL       1     21       BY SIMILARITY.
FT   CHAIN       22    678       OVIDUCT-SPECIFIC GLYCOPROTEIN.
FT   CARBOHYD   402    402       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   441    441       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   580    580       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   596    596       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   648    648       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT   477    477       M -> T (IN REF. 2).
FT   CONFLICT   511    511       S -> P (IN REF. 2).
FT   CONFLICT   514    514       Y -> H (IN REF. 2).
FT   CONFLICT   676    676       E -> Q (IN REF. 2).
SQ   SEQUENCE   678 AA;  75421 MW;  245F2CEDCE92768B CRC64;

  Query Match             55.4%;  Score 36;  DB 1;  Length 678;
  Best Local Similarity   66.7%;  Pred. No. 39;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 PTFGRSFTL 9
              ||:|||:| |
Db        261 PTYGRTFRL 269


RESULT 15
ZDH5_HUMAN
ID   ZDH5_HUMAN     STANDARD;      PRT:   715 AA.
AC   Q9C0B5; Q8TAK8; Q9H923; Q9UFI7;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Zinc finger DHHC domain containing protein 5 (Zinc finger protein
DE   375).
GN   ZDHHC5 OR ZNF375 OR KIAA1748.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RX   MEDLINE=21082932; PubMed=11214970;
RA   Nagase T., Kikuno R., Hattori A., Kondo Y., Okumura K., Ohara O.;
RT   "Prediction of the coding sequences of unidentified human genes. XIX.
RT   The complete sequences of 100 new cDNA clones from brain which code
RT   for large proteins in vitro.";
RL   DNA Res. 7:347-355(2000).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RA   Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA   Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
RA   Wagatsuma M., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA   Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA   Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA   Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,
RA   Ninomiya K., Iwayanagi T.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM 2).
RC   TISSUE=Uterus;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
```

```
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [4]
RP   SEQUENCE OF 508-715 FROM N.A. (ISOFORM 1/2).
RC   TISSUE=Uterus;
RA   Duesterhoeft A., Lauber J., Mewes H.-W., Rossenkuhler J., Wiemann S.;
RL   Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1;
CC         IsoId=Q9C0B5-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=Q9C0B5-2; Sequence=VSP_006935;
CC         Note=No experimental confirmation available;
CC   -!- SIMILARITY: Contains 1 DHHC-type zinc finger.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AB051535; BAB21839.1; ALT_INIT.
DR   EMBL; AK023130; BAB14420.1; -.
DR   EMBL; BC026967; AAH26967.1; -.
DR   EMBL; AL117662; CAB56033.1; -.
DR   PIR; T17343; T17343.
DR   Genew; HGNC:18472; ZDHHC5.
DR   InterPro; IPR001594; Znf_DHHC.
DR   Pfam; PF01529; zf-DHHC; 1.
DR   ProDom; PD003041; Znf_DHHC; 1.
DR   PROSITE; PS50216; ZF_DHHC; 1.
KW   Transmembrane; Zinc-finger; Alternative splicing.
FT   TRANSMEM      14     34       POTENTIAL.
FT   TRANSMEM      39     59       POTENTIAL.
FT   TRANSMEM     149    169       POTENTIAL.
FT   TRANSMEM     192    212       POTENTIAL.
FT   ZN_FING      104    154       DHHC-TYPE.
FT   VARSPLIC       1     53       Missing (in isoform 2).
FT                                 /FTId=VSP_006935.
FT   CONFLICT     508    509       QQ -> TR (IN REF. 4).
SQ   SEQUENCE   715 AA;  77544 MW;  9E4FB0C9AC8EFE28 CRC64;

  Query Match             55.4%;  Score 36;  DB 1;  Length 715;
  Best Local Similarity   85.7%;  Pred. No. 42;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PTFCRSF 7
              ||:|:||
Db        410 PTFGKSF 416


Search completed: September 24, 2003, 13:42:27
Job time : 9.17978 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 24, 2003, 13:33:45 ; Search time 42.2135 Seconds
(without alignments)
79.469 Million cell updates/sec

Title: US-09-744-282-5
Perfect score: 65
Sequence: 1 PTFGRSFTLASSE 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	65	100.0	383	4	Q96HI7	Q96hi7 homo sapien
2	65	100.0	383	4	Q8IVA4	Q8iva4 homo sapien
3	62	95.4	352	11	Q9WTV1	Q9wtv1 rattus norv
4	62	95.4	381	11	Q99J84	Q99j84 mus musculu
5	62	95.4	389	11	Q8BKL8	Q8bkl8 mus musculu
6	61	93.8	383	6	Q6SPQ0	Q6spq0 capra hircu
7	61	93.8	383	6	Q29411	Q29411 sus scrofa
8	58	89.2	332	6	O18949	O18949 bos taurus
9	57	87.7	387	4	Q9H3V8	Q9h3v8 homo sapien
10	57	87.7	396	11	Q9D7Q1	Q9d7q1 mus musculu
11	57	87.7	466	4	Q13231	Q13231 homo sapien
12	54	83.1	390	4	Q96F97	Q96f97 homo sapien
13	51	78.5	983	5	Q9V2V2	Q9v2v2 drosophila
14	51	78.5	1013	5	Q960M0	Q960m0 drosophila
15	49	75.4	268	5	Q8MSI4	Q8msi4 drosophila
16	49	75.4	460	5	Q9W2M7	Q9w2m7 drosophila
17	47	72.3	4498	5	Q9W2Z3	Q9w2z3 drosophila
18	45	69.2	525	5	O44079	O44079 anopheles g
19	44	67.7	431	5	Q8ISH5	Q8ish5 araneus ven
20	43	66.2	553	5	P91731	P91731 hyphantria
21	43	66.2	929	5	Q8MY79	Q8my79 haemaphysal
22	43	66.2	1080	5	Q9BLI6	Q9bli6 bombyx mori
23	42	64.6	2838	5	Q8MP05	Q8mp05 tenebrio mo
24	41	63.1	276	5	Q9VIZ4	Q9viz4 drosophila
25	41	63.1	432	5	Q9W3V1	Q9w3v1 drosophila
26	41	63.1	574	5	O17411	O17411 aedes aegyp
27	41	63.1	595	5	Q9VFR3	Q9vfr3 drosophila
28	40	61.5	358	16	Q8CX87	Q8cx87 oceanobacil
29	40	61.5	415	16	Q9AAG2	Q9aag2 caulobacter
30	40	61.5	543	5	Q9GR93	Q9gr93 bombyx mori
31	40	61.5	543	5	Q9GV05	Q9gv05 bombyx mori
32	40	61.5	544	5	Q9GQC4	Q9gqc4 bombyx mori
33	40	61.5	552	5	Q9GV44	Q9gv44 spodoptera
34	40	61.5	557	5	Q8MTK0	Q8mtk0 choristonei
35	40	61.5	565	5	P90710	P90710 bombyx mori
36	40	61.5	565	5	Q9GPG9	Q9gpg9 bombyx mand
37	40	61.5	586	5	Q8WR52	Q8wr52 bombyx mori
38	40	61.5	1635	5	O17412	O17412 aedes aegyp
39	39	60.0	113	13	Q91153	Q91153 notophthalm
40	39	60.0	280	13	Q91152	Q91152 notophthalm
41	39	60.0	329	4	Q96DI0	Q96di0 homo sapien
42	39	60.0	467	5	O15993	O15993 penaeus jap
43	39	60.0	467	5	Q8ITU3	Q8itu3 penaeus van
44	39	60.0	483	5	Q23737	Q23737 chelonus sp
45	39	60.0	497	5	Q25615	Q25615 onchocerca

ALIGNMENTS

RESULT 1
Q96HI7
ID Q96HI7 PRELIMINARY; PRT; 383 AA.
AC Q96HI7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Similar to chitinase 3-like 1 (Cartilage glycoprotein-39).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC008568; AAH08568.1; -.
DR InterPro; IPR001223; Glyco_hydro_18.
DR InterPro; IPR005829; Sug_transporter.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR SMART; SM00636; Glyco_18; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
SQ SEQUENCE 383 AA; 42625 MW; 6D9EF133BDC7C2D1 CRC64;

Query Match 100.0%; Score 65; DB 4; Length 383;
Best Local Similarity 100.0%; Pred. No. 0.00037;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PTFGRSFTLASSE 13
|||||||||||||
Db 259 PTFGRSFTLASSE 271

RESULT 2
Q8IVA4
ID Q8IVA4 PRELIMINARY; PRT; 383 AA.
AC Q8IVA4;

```
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Similar to chitinase 3-like 1 (Cartilage glycoprotein-39).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RA   Strausberg R.;
RL   Submitted (OCT-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC038354; AAH38354.1; -.
SQ   SEQUENCE   383 AA;  42526 MW;  76345433BDC7C6C0 CRC64;

  Query Match           100.0%;  Score 65;  DB 4;  Length 383;
  Best Local Similarity  100.0%;  Pred. No. 0.00037;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PTFGRSFTLASSE 13
              |||||||||||||
Db        259 PTFGRSFTLASSE 271


RESULT 3
Q9WTV1
ID   Q9WTV1      PRELIMINARY;      PRT;   352 AA.
AC   Q9WTV1;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Glycoprotein-39 (Fragment).
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Lewis;
RA   Wendling U., Boots A.M.H., van Eden W.;
RT   "Cloning of the rat homologue of Human Cartilage glycoprotein-39 a
RT   potential autoantigen in arthritis.";
RL   Submitted (APR-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF062038; AAD22610.1; -.
DR   HSSP; P07254; 1CTN.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   ProDom; PD000471; Glyco_hydro_18; 1.
DR   SMART; SM00636; Glyco_18; 1.
FT   NON_TER       1      1
FT   NON_TER     352    352
SQ   SEQUENCE   352 AA;  39391 MW;  CBDE991610AC936C CRC64;

  Query Match           95.4%;  Score 62;  DB 11;  Length 352;
  Best Local Similarity  92.3%;  Pred. No. 0.0013;
  Matches   12;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PTFGRSFTLASSE 13
              |||:|||||||||
Db        238 PTFGKSFTLASSE 250


RESULT 4
Q99J84
ID   Q99J84      PRELIMINARY;      PRT;   381 AA.
AC   Q99J84;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Similar to chitinase 3-like 1 (Cartilage glycoprotein-39).
GN   CHI3L1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Strausberg R.;
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Strausberg R.;
RL   Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Strausberg R.;
RL   Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC005611; AAH05611.1; -.
DR   EMBL; BC003780; AAH03780.1; -.
DR   EMBL; BC004734; AAH04734.1; -.
DR   MGD; MGI:1340899; Chi3l1.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   ProDom; PD000471; Glyco_hydro_18; 1.
DR   SMART; SM00636; Glyco_18; 1.
SQ   SEQUENCE   381 AA;  42979 MW;  EF6588C5AE9D4450 CRC64;

  Query Match           95.4%;  Score 62;  DB 11;  Length 381;
  Best Local Similarity  92.3%;  Pred. No. 0.0014;
  Matches   12;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PTFGRSFTLASSE 13
              |||:|||||||||
Db        260 PTFGKSFTLASSE 272


RESULT 5
Q8BKL8
ID   Q8BKL8      PRELIMINARY;      PRT;   389 AA.
AC   Q8BKL8;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Similar to chitinase 3-like 1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Dorsal root ganglion;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
DR   EMBL; AK051475; BAC34654.1; -.
SQ   SEQUENCE   389 AA;  43893 MW;  9E7D66063B233834 CRC64;

  Query Match           95.4%;  Score 62;  DB 11;  Length 389;
  Best Local Similarity  92.3%;  Pred. No. 0.0014;
  Matches   12;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PTFGRSFTLASSE 13
              |||:|||||||||
Db        268 PTFGKSFTLASSE 280


RESULT 6
Q8SPQ0
ID   Q8SPQ0      PRELIMINARY;      PRT;   383 AA.
AC   Q8SPQ0;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
```

```
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   BP40 precursor.
OS   Capra hircus (Goat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Caprinae; Capra.
OX   NCBI_TaxID=9925;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Mammary gland;
RA   Paramasivam M., Saravanan K., Sharma S., Mohanty A.K., Yadav S.,
RA   Srinivasan A., Singh T.P.;
RT   "Goat mammary gland mRNA for a novel protein, BP40, complete coding
RT   region.";
RL   Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY081150; AAL87607.1; -.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   ProDom; PD000471; Glyco_hydro_18; 1.
DR   SMART; SM00636; Glyco_18; 1.
KW   Signal.
FT   SIGNAL       1     16       POTENTIAL.
SQ   SEQUENCE   383 AA;  42893 MW;  17655BD4BE4E9E5F CRC64;

  Query Match             93.8%;  Score 61;  DB 6;  Length 383;
  Best Local Similarity   92.3%;  Pred. No. 0.0022;
  Matches   12;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PTFGRSFTLASSE 13
              |:||||:|||||:
Db        259 PTFGRSFTLASSK 271


RESULT 7
Q29411
ID   Q29411      PRELIMINARY;      PRT;   383 AA.
AC   Q29411;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   38 kDa heparin-binding glycoprotein.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Smooth muscle;
RX   MEDLINE=95286589; PubMed=7768902;
RA   Shackelton L.M., Mann D.M., Millis A.J.;
RT   "Identification of a 38-kDa heparin-binding glycoprotein (gp38k) in
RT   differentiating vascular smooth muscle cells as a member of a group of
RT   proteins associated with tissue remodeling.";
RL   J. Biol. Chem. 270:13076-13083(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Smooth muscle;
RA   Shackelton L.M., Mann D.M., Millis A.J.T.;
RL   Submitted (JAN-1995) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U19900; AAA86482.1; -.
DR   EMBL; Z47803; CAA87764.1; -.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   ProDom; PD000471; Glyco_hydro_18; 1.
DR   SMART; SM00636; Glyco_18; 1.
KW   Heparin-binding.
SQ   SEQUENCE   383 AA;  42443 MW;  3D1039F49910BDC4 CRC64;

  Query Match             93.8%;  Score 61;  DB 6;  Length 383;
  Best Local Similarity   92.3%;  Pred. No. 0.0022;
  Matches   12;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PTFGRSFTLASSE 13
              |:||||:|||||:
Db        259 PTFGRSFTLASSK 271


RESULT 8
Q18949
ID   Q18949      PRELIMINARY;      PRT;   332 AA.
AC   Q18949;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Chitinase-like protein 1 (Fragment).
GN   CLP-1.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Recklies A.D., White C.;
RT   "Expression of chitinase-like protein 1 (CLP-1) in bovine
RT   chondrocytes.";
RL   Submitted (JUN-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF011373; AAB64304.1; -.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   ProDom; PD000471; Glyco_hydro_18; 1.
DR   SMART; SM00636; Glyco_18; 1.
FT   NON_TER       1       1
FT   NON_TER     332     332
SQ   SEQUENCE   332 AA;  37434 MW;  03F163326486408F CRC64;

  Query Match             89.2%;  Score 58;  DB 6;  Length 332;
  Best Local Similarity   84.6%;  Pred. No. 0.0072;
  Matches   11;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PTFGRSFTLASSE 13
              |||||:|||||:
Db        235 PTFGRSYTLASSK 247


RESULT 9
Q9H3V8
ID   Q9H3V8      PRELIMINARY;      PRT;   387 AA.
AC   Q9H3V8;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Chitotriosidase precursor.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96064695; PubMed=7592832;
RA   Boot R.G., Renkema G.H., Strijland A., van Zonneveld A.J., Aerts J.M.;
RT   "Cloning of a cDNA encoding chitotriosidase, a human chitinase
RT   produced by macrophages.";
RL   J. Biol. Chem. 270:26252-26256(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98421482; PubMed=9748235;
RA   Boot R.G., Renkema G.H., Verhoek M., Strijland A., Bliek J.,
RA   de Meulemeester T.M., Mannens M.M., Aerts J.M.;
RT   "The human chitotriosidase gene. Nature of inherited enzyme
RT   deficiency.";
RL   J. Biol. Chem. 273:25680-25685(1998).
DR   EMBL; U62662; AAG10644.1; -.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   InterPro; IPR001579; Glyco_hydro_18/2.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
```

```
DR   ProDom; PD000471; Glyco_hydro_18; 1.
DR   SMART; SM00636; Glyco_18; 1.
DR   PROSITE; PS01095; CHITINASE_18; 1.
KW   Signal.
FT   SIGNAL        1     21       POTENTIAL.
FT   CHAIN        22    387       CHITOTRIOSIDASE.
SQ   SEQUENCE   387 AA;  43133 MW;  03A272B8BC5E0D71 CRC64;

  Query Match             87.7%;  Score 57;  DB 4;  Length 387;
  Best Local Similarity   91.7%;  Pred. No. 0.013;
  Matches   11;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PTFGRSFTLASS 12
              ||:|||||||||
Db        265 PTYGRSFTLASS 276


RESULT 10
Q9D7Q1
ID   Q9D7Q1      PRELIMINARY;      PRT;   396 AA.
AC   Q9D7Q1;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   2300002L19Rik protein.
GN   2300002L19RIK.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Tongue;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
DR   EMBL; AK009012; BAB26025.1; -.
DR   MGD; MGI:1919134; 2300002L19Rik.
DR   InterPro; IPR000877; 2S_Globulin.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   InterPro; IPR001579; Glyco_hydro_18/2.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   PRINTS; PR00551; 2SGLOBULIN.
DR   ProDom; PD000471; Glyco_hydro_18; 1.
DR   SMART; SM00636; Glyco_18; 1.
DR   PROSITE; PS01095; CHITINASE_18; 1.
SQ   SEQUENCE   396 AA;  43413 MW;  32BB674B61AF7123 CRC64;

  Query Match             87.7%;  Score 57;  DB 11;  Length 396;
  Best Local Similarity   91.7%;  Pred. No. 0.014;
  Matches   11;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PTFGRSFTLASS 12
              ||:|||||||||
Db        265 PTYGRSFTLASS 276


RESULT 11
Q13231
ID   Q13231      PRELIMINARY;      PRT;   466 AA.
AC   Q13231;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Chitotriosidase precursor.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96064695; PubMed=7592832;
RA   Boot R.G., Renkema G.H., Strijland A., van Zonneveld A.J., Aerts J.M.;
RT   "Cloning of a cDNA encoding chitotriosidase, a human chitinase
RT   produced by macrophages.";
RL   J. Biol. Chem. 270:26252-26256(1995).
DR   EMBL; U29615; AAC50246.1; -.
DR   Genew; HGNC:1936; CHIT1.
DR   InterPro; IPR002557; Chitin_bind_PerA.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   InterPro; IPR001579; Glyco_hydro_18/2.
DR   Pfam; PF01607; CBM_14; 1.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   ProDom; PD000471; Glyco_hydro_18; 1.
DR   SMART; SM00494; ChtBD2; 1.
DR   SMART; SM00636; Glyco_18; 1.
DR   PROSITE; PS01095; CHITINASE_18; 1.
KW   Signal.
FT   SIGNAL        1     21       POTENTIAL.
FT   CHAIN        22    466       CHITOTRIOSIDASE.
SQ   SEQUENCE   466 AA;  51681 MW;  B4312D1E885E386D CRC64;

  Query Match             87.7%;  Score 57;  DB 4;  Length 466;
  Best Local Similarity   91.7%;  Pred. No. 0.016;
  Matches   11;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 PTFGRSFTLASS 12
              ||:|||||||||
Db        265 PTYGRSFTLASS 276


RESULT 12
Q96F97
ID   Q96F97      PRELIMINARY;      PRT;   390 AA.
AC   Q96F97;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Chitinase 3-like 2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RA   Strausberg R.;
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC011460; AAH11460.1; -.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   ProDom; PD000471; Glyco_hydro_18; 1.
DR   SMART; SM00636; Glyco_18; 1.
SQ   SEQUENCE   390 AA;  43529 MW;  6801E4AF721AD63F CRC64;

  Query Match             83.1%;  Score 54;  DB 4;  Length 390;
  Best Local Similarity   76.9%;  Pred. No. 0.051;
  Matches   10;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;
```

```
Qy          1 PTFGRSFTLASSE 13
              ||:| ||||||:.
Db        267 PTYGHSFTLASAE 279


RESULT 13
Q9VZV2
ID   Q9VZV2      PRELIMINARY;      PRT;   983 AA.
AC   Q9VZV2;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   CG1869 protein.
GN   CG1869.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.C.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
RA   Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA   Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
RA   Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA   Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnok D., Farfan D.,
RA   Ferriera S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA   Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA   Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA   McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA   Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA   Phouanenavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA   Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
RA   Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT   "Sequencing of Drosophila melanogaster genome.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA   Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
RA   Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
RA   Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA   Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA   Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA   Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT   "Annotation of Drosophila melanogaster genome.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE FROM N.A.
RA   FlyBase;
RL   Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AE003477; AAF47714.2; -.
DR   FlyBase; FBgn0035398; CG1869.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   InterPro; IPR001579; Glyco_hydro_18/2.
DR   Pfam; PF00704; Glyco_hydro_18; 2.
DR   ProDom; PD000471; Glyco_hydro_18; 2.
DR   PROSITE; PS01095; CHITINASE_18; 2.
SQ   SEQUENCE   983 AA;  109729 MW;  59A8ED4A17B44051 CRC64;

  Query Match             78.5%;  Score 51;  DB 5;  Length 983;
  Best Local Similarity   69.2%;  Pred. No. 0.52;
  Matches    9;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 PTFGRSFTLASSE 13
              ||:|||||||: :
Db        368 PTYGRSFTLANPD 380


RESULT 14
Q960M0
ID   Q960M0      PRELIMINARY;      PRT;  1013 AA.
AC   Q960M0;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   LD45559p.
GN   CG1869.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RA   Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA   Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA   Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA   Nunoo J., Pacleb J., Paragas V., Park S., Phouanenavong S., Wan K.,
RA   Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL   Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY051988; AAK93412.1; -.
DR   FlyBase; FBgn0035398; CG1869.
DR   InterPro; IPR002557; Chitin_bind_PerA.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   InterPro; IPR001579; Glyco_hydro_18/2.
DR   Pfam; PF01607; CBM_14; 1.
DR   Pfam; PF00704; Glyco_hydro_18; 2.
```

```
DR   ProDom; PD000471; Glyco_hydro_18; 2.
DR   SMART; SM00494; ChtBD2; 1.
DR   SMART; SM00636; Glyco_18; 2.
DR   PROSITE; PS01095; CHITINASE_18; 2.
SQ   SEQUENCE   1013 AA;  113285 MW;  266981AD1685EC02 CRC64;

  Query Match             75.5%;  Score 51;  DB 5;  Length 1013;
  Best Local Similarity   69.2%;  Pred. No. 0.54;
  Matches    9;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 PTFGRSFTLASSE 13
              |:|:||||:|: :
Db        368 PTYGRSFTLANPD 380


RESULT 15
Q8MSI4
ID   Q8MSI4      PRELIMINARY;      PRT;   268 AA.
AC   Q8MSI4;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   GH13872p.
GN   CG9357.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RA   Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA   Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA   George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA   Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA   Patel S., Phouanenavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA   Celniker S.;
RL   Submitted (JUN-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY118784; AAM50644.1; -.
DR   FlyBase; FBgn0034580; CG9357.
DR   InterPro; IPR002557; Chitin_bind_PerA.
DR   InterPro; IPR001179; PKBP_PPIase.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF01607; CBM_14; 1.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   ProDom; PD000471; Glyco_hydro_18; 1.
DR   SMART; SM00494; ChtBD2; 1.
DR   SMART; SM00636; Glyco_18; 1.
DR   PROSITE; PS00453; FKBP_PPIASE_1; 1.
SQ   SEQUENCE   268 AA;  29678 MW;  5DEBDC9903CB2D1E CRC64;

  Query Match             75.4%;  Score 49;  DB 5;  Length 268;
  Best Local Similarity   69.2%;  Pred. No. 0.31;
  Matches    9;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 PTFGRSFTLASSE 13
              | :|:|||||:|:|
Db         70 PFYGRSFTLATAE 82


Search completed: September 24, 2003, 13:47:25
Job time : 43.2135 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 24, 2003, 13:25:57 ; Search time 52.1461 Seconds
(without alignments)
39.570 Million cell updates/sec

Title: US-09-744-282-6
Perfect score: 62
Sequence: 1 RSFTLASSEIGVG 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000001000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	62	100.0	13	17	AAR95076	Peptide contg. hC
2	62	100.0	13	19	AAW61597	Human Cartilage gp
3	62	100.0	13	21	AAY69231	Peptide fragment c
4	62	100.0	13	21	AAY77401	Human cartilage gl
5	62	100.0	13	22	AAU03045	Human cartilage gp
6	62	100.0	13	22	AAU03048	Human cartilage gp
7	62	100.0	13	22	AAU03059	Human cartilage gp
8	62	100.0	13	22	AAU03065	Human cartilage gp
9	62	100.0	13	22	AAU03066	Human cartilage gp
10	62	100.0	13	22	AAU03067	Human cartilage gp
11	62	100.0	13	23	AAE20207	Human cartilage gl
12	62	100.0	13	23	AAE20214	Human cartilage gl
13	62	100.0	13	23	AAE20215	Human cartilage gl
14	62	100.0	13	23	AAE20216	Human cartilage gl
15	62	100.0	13	23	AAE20220	Human cartilage gl
16	62	100.0	15	19	AAW61598	Human Cartilage gp
17	62	100.0	16	18	AAW45140	Human cartilage gl
18	62	100.0	16	19	AAW61599	Human Cartilage gp
19	62	100.0	105	22	AAU02074	Synthetic human ta
20	62	100.0	199	22	ABB56452	Human single chain
21	62	100.0	199	22	ABB56456	Human single chain
22	62	100.0	208	22	ABB56475	Human single chain
23	62	100.0	209	22	ABB56454	Human single chain
24	62	100.0	210	22	ABB56442	Human single chain
25	62	100.0	211	22	ABB56441	Human single chain
26	62	100.0	211	22	ABB56455	Human single chain
27	62	100.0	212	22	ABB56439	Human single chain
28	62	100.0	212	22	ABB56440	Human single chain
29	62	100.0	212	22	ABB56477	Human single chain
30	62	100.0	213	22	ABB56443	Human single chain
31	62	100.0	213	22	ABB56445	Human single chain
32	62	100.0	213	22	ABB56446	Human single chain
33	62	100.0	213	22	ABB56453	Human single chain
34	62	100.0	214	22	ABB56444	Human single chain
35	62	100.0	216	22	ABB56447	Human single chain
36	62	100.0	216	22	ABB56448	Human single chain
37	62	100.0	220	22	ABB56449	Human single chain
38	62	100.0	220	22	ABB56450	Human single chain
39	62	100.0	220	22	AAU02078	Synthetic human mu
40	62	100.0	281	22	AAU02079	Synthetic multi-ta
41	62	100.0	350	22	AAU02076	Synthetic human mu
42	62	100.0	383	23	ABC96297	Human ovarian canc
43	62	100.0	383	24	ABU55651	Lung cancer-associ
44	62	100.0	459	22	AAU02077	Synthetic multi-ta
45	58	93.5	13	22	AAU03054	Human cartilage gp

ALIGNMENTS

RESULT 1
AAR95076
ID AAR95076 standard; peptide; 13 AA.
XX
AC AAR95076;
XX
DT 17-DEC-1996 (first entry)
XX
DE Peptide contg. hC gp-39 autoantigen derived peptide (II).
XX
KW Human; cartilage; glycoprotein 39; hC gp-39; autoantigen;
KW induction; T-cell; tolerance; mediated; destruction;
KW rheumatoid arthritis; disease model; diagnosis.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 3..11
FT /note= "HC gp-39 peptide (II)"
XX
PN WO9613517-A1.
XX
PD 09-MAY-1996.
XX
PF 25-OCT-1995; 95WO-EP04201.
XX
PR 07-APR-1995; 95EP-0200886.
PR 27-OCT-1994; 94EP-0203128.
XX
PA (AIKU) AKZO NOBEL NV.
XX
PI Boots AMH, Verheijden GPM;

```
XX
DR   WPI; 1999-239447/24.
XX
PT   Peptide(s) derived from human cartilage gp-39 autoantigen - used
PT   to induce specific T cell tolerance, esp. for treatment of
PT   rheumatoid arthritis, also useful in producing arthritic animal
PT   models
XX
PS   Claim 2; Page 28; 37pp; English.
XX
CC   The present peptide, contg. a peptide derived from the human
CC   cartilage glycoprotein 39 (hC gp-39) autoantigen, can be used to
CC   induce specific T-cell tolerance to hC gp-39 in a patient with
CC   T-cell mediated cartilage destruction, esp. rheumatoid arthritis
CC   (RA). It may also be used to induce RA in mice, etc. (human disease
CC   models), and to diagnose the presence of activated T-cells
CC   autoreactive with hC gp-39, i.e. to identify patients suitable for
CC   tolerising treatment.
CC   Therapeutically the peptide is given in a dose of 0.1-100 microg/kg
CC   by injection, or 10-50 microg/kg to produce a disease model. The
CC   peptide has a specific tolerising effect on autoreactive T-cells,
CC   but leaves other components of the immune system intact, and
CC   therefore should have few side effects.
XX
SQ   Sequence   13 AA;

  Query Match             100.0%;  Score 62;  DB 17;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 6e-05;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RSFTLASSETGVG 13
              |||||||||||||
Db          1 RSFTLASSETGVG 13


RESULT 2
AAW61597
ID   AAW61597 standard; protein; 13 AA.
XX
AC   AAW61597;
XX
DT   27-OCT-1998  (first entry)
XX
DE   Human Cartilage gp-39 epitope (1).
XX
KW   Human; cartilage gp-39; monoclonal antibody; cell surface antigen;
KW   rheumatoid arthritis; T-cell receptor.
XX
OS   Homo sapiens.
XX
PN   EP856520-A1.
XX
PD   05-AUG-1998.
XX
PF   02-DEC-1997;   97EP-0203769.
XX
PR   27-JUN-1997;   97EP-0201972.
PR   06-DEC-1996;   96EP-0203465.
XX
PA   (ALKU ) AKZO NOBEL NV.
PA   (ALKU ) AKZO NV.
XX
PI   Steenbakkers PGA;
XX
DR   WPI; 1998-400866/35.
XX
PT   Preparation of monoclonal antibodies against cell surface antigens
PT   used as diagnostic agents and to treat rheumatoid arthritis -
PT   comprises immunising, then enriching, enriching and cloning B cells
PT   and then selecting and cloning hybridoma
XX
PS   Examples; Page 5; 23pp; English.
XX
CC   The human cartilage gp-39 peptides AAW61597-W61599 were used in the
CC   preparation of a monoclonal antibody against a cell surface antigen.
CC   The antibody can be used for treatment of rheumatoid arthritis.  It
CC   is also useful against cell surface antigens and antibodies reactive
CC   with the clonotypic structure of T-cell receptors are used as diagnostic
CC   reagents.  The process allows antibodies to be raised against
CC   antigens which are present in very low amounts or which are
CC   conformationally unstable without the need to screen very large numbers
CC   of hybridomas and without requiring high purity Antigen for enrichment.
XX
SQ   Sequence   13 AA;

  Query Match             100.0%;  Score 62;  DB 19;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 6e-05;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RSFTLASSETGVG 13
              |||||||||||||
Db          1 RSFTLASSETGVG 13


RESULT 3
AAY69231
ID   AAY69231 standard; peptide; 13 AA.
XX
AC   AAY69231;
XX
DT   30-MAY-2000  (first entry)
XX
DE   Peptide fragment of human cartilage gp-39 (HC gp-39) protein.
XX
KW   Human cartilage gp-39 protein; HC gp-39 protein; T-cell activator;
KW   inflammatory disease; lymphocyte; antigen; bystander suppression;
KW   antigen-specific T-cell tolerance; autoimmune disorder;
KW   inflammatory disease; rheumatoid arthritis;
KW   autoimmune inflammatory disease; autoantigen.
XX
OS   Homo sapiens.
XX
PN   WO200004917-A2.
XX
PD   03-FEB-2000.
XX
PF   19-JUL-1999;   99WO-EP05331.
XX
PR   23-JUL-1998;   98EP-0202471.
XX
PA   (ALKU ) AKZO NOBEL NV.
XX
PI   Miltenburg AMM,  Boots AMH;
XX
DR   WPI; 2000-182537/16.
XX
PT   Use of cartilage protein or peptides for treating inflammatory diseases
PT   especially rheumatoid arthritis -
XX
PS   Claim 2; Page 17; 29pp; English.
XX
CC   AAY69226-33 represent fragments of the human cartilage gp-39 (HC gp-39)
CC   protein. The peptides are T-cell activators. The peptides are used for
CC   preventing inflammatory diseases, by modulating lymphocytes reactive
CC   to antigens other than HC gp-39, present in the same tissue of
CC   HC gp-39. By the induction of antigen-specific T-cell tolerance,
CC   autoimmune disorders can be treated by bystander suppression. The
CC   HC gp-39 peptides are useful for preventing inflammatory diseases
CC   especially rheumatoid arthritis. The fragments may also be used for
CC   treating autoimmune inflammatory diseases. The HC gp-39 peptides are
CC   especially beneficial in autoimmune conditions in which the
CC   autoantigens associated with the condition are unknown.
XX
SQ   Sequence   13 AA;

  Query Match             100.0%;  Score 62;  DB 21;  Length 13;
```

```
  Best Local Similarity   100.0%;  Pred. No. 6e-05;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RSFTLASSETGVG 13
              |||||||||||||
Db          1 RSFTLASSETGVG 13


RESULT 4
AAY77401
ID   AAY77401 standard; peptide; 13 AA.
XX
AC   AAY77401;
XX
DT   22-MAY-2000  (first entry)
XX
DE   Human cartilage glycoprotein-39 (HC gp-39), residues 263-275.
XX
KW   Human cartilage glycoprotein-39; HC gp-39; YKL-40; epitope;
KW   autoreactive T-cell; tolerisation; desensitisation; autoantigen;
KW   rheumatoid arthritis; human chondrocyte protein; YKL-39;
KW   bystander suppression; autoimmune disorder; Graves' disease;
KW   primary glomerulonephritis; inflammatory bowel disease; diagnosis.
XX
OS   Homo sapiens.
XX
PN   WO200005254-A2.
XX
PD   03-FEB-2000.
XX
PF   16-JUL-1999;   99WO-EP05050.
XX
PR   23-JUL-1998;   98EP-0202470.
XX
PA   (ALKU ) AKZO NOBEL NV.
XX
PI   Verheijden GFM,  Boots AMH;
XX
DR   WPI; 2000-182644/16.
XX
PT   New T-cell epitope peptides useful for inducing specific T-cell
PT   tolerance to an autoantigen for treating autoimmune disorders,
PT   especially rheumatoid arthritis
XX
PS   Example 1; Page 10; 26pp; English.
XX
CC   The invention relates to peptides constituting amino acids 268-276 and
CC   amino acids 266-278 of human chondrocyte protein YKL-39
CC   (AAY77399-Y77400, respectively). The peptides of the invention may be
CC   used in pharmaceutical compositions for the induction of specific T-cell
CC   tolerance (desensitisation) to an autoantigen in patients
CC   suffering from autoimmune disorders, especially rheumatoid arthritis. A
CC   frequently found autoantigen in rheumatoid arthritis sufferers is human
CC   cartilage glycoprotein-39 (HC gp-39, also known as YKL-40). Several
CC   epitopes of HC gp-39 are recognised by peripheral blood T-cells from
CC   rheumatoid arthritis patients; one such epitope, HC gp-39 residues
CC   263-275, was recognised by T-cells from 8 out of 18 patients tested.
CC   YKL-39 shares significant sequence identity to hc gp-39. YKL-39 residues
CC   266-278 (AAY77400) constitutes a mimicry epitope of HC gp-39 (263-275),
CC   and may be used for desensitisation of T-cells with reactivity to HC
CC   gp-39 (263-275) in rheumatoid arthritis patients. Rheumatoid arthritis
CC   may be treated by the systemic administration of compositions comprising
CC   the peptides of the invention, optionally in combination with existing
CC   treatments. The peptides may also be used for modulating lymphocytes
CC   that are reactive to autoantigens other than YKL-39 and HC gp-39 but
CC   which are present in the same tissue as these autoantigens. By the
CC   induction of antigen-specific T-cell tolerance, autoimmune disorders can
CC   be treated by bystander suppression. The conditions that may be treated
CC   via this approach include Graves' disease, primary glomerulonephritis
CC   and inflammatory bowel disease. The peptides, along with a detection
CC   agent, may additionally be used in a diagnostic composition. The peptides
CC   of the invention have a specific effect on autoreactive T-cells whilst
CC   leaving other components of the immune system intact. Therefore, the
CC   peptides should be safe, having no side effects. The present sequence
CC   represents residues 263-275 of human cartilage glycoprotein-39 (HC gp-39)
CC   which was used in an exemplification of the present invention.
XX
SQ   Sequence   13 AA;

  Query Match             100.0%;  Score 62;  DB 21;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 6e-05;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RSFTLASSETGVG 13
              |||||||||||||
Db          1 RSFTLASSETGVG 13


RESULT 5
AAU03045
ID   AAU03045 standard; peptide; 13 AA.
XX
AC   AAU03045;
XX
DT   12-SEP-2001  (first entry)
XX
DE   Human cartilage gp-39 autoimmune peptide #1.
XX
KW   Human; autoimmune disease; epitope; antigen; rheumatoid arthritis;
KW   immunosuppression.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "Deaminated"
XX
PN   WO200129081-A1.
XX
PD   26-APR-2001.
XX
PF   12-OCT-2000; 2000WO-EP10230.
XX
PR   18-OCT-1999;   99EP-0203427.
XX
PA   (ALKU ) AKZO NOBEL NV.
XX
PI   Van Staveren CJ,  Timmers CM,  Van Galen PJM,  Knegtel RMA,  Boots AMH;
PI   Miltenburg AMM;
XX
DR   WPI; 2001-367293/38.
XX
PT   Modified peptides derived from human cartilage gp-39 protein epitope
PT   are useful in immunotherapy, especially in the treatment of arthritis
PT   -
XX
PS   Claim 1; Page 36; 52pp; English.
XX
CC   The sequence represents Human cartilage gp-39 protein epitope (amino
CC   acids 263-275) which is used as the template for making modified
CC   peptides for use in immunotherapy. The peptides are useful for inducing
CC   tolerance induction in patients suffering from autoimmune diseases such
CC   as rheumatoid arthritis. The peptides provide antigen-specific, non-toxic
CC   immunosuppression therapy as a very attractive alternative to
CC   non-specific immunosuppression.
XX
SQ   Sequence   13 AA;

  Query Match             100.0%;  Score 62;  DB 22;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 6e-05;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RSFTLASSETGVG 13
              |||||||||||||
Db          1 RSFTLASSETGVG 13
```

```
RESULT 6
AAU03048
ID   AAU03048 standard; peptide; 13 AA.
XX
AC   AAU03048;
XX
DT   12-SEP-2001  (first entry)
XX
DE   Human cartilage gp-39 autoimmune peptide #4.
XX
KW   Human; autoimmune disease; epitope; antigen; rheumatoid arthritis;
KW   immunosuppression.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "Connected to CH3-(OCH2CH3)3-OCH2C(O)"
FT   Modified-site   13
FT                   /note= "C-terminus is amidated"
XX
PN   WO200129081-A1.
XX
PD   26-APR-2001.
XX
PF   12-OCT-2000; 2000WO-EP10230.
XX
PR   18-OCT-1999;   99EP-0203427.
XX
PA   (ALKU ) AKZO NOBEL NV.
XX
PI   Van Staveren CJ,  Timmers CM,  Van Galen PJM,  Knegtel RMA,  Boots AMH;
PI   Miltenburg AMM;
XX
DR   WPI; 2001-367293/38.
XX
PT   Modified peptides derived from human cartilage gp-39 protein epitope
PT   are useful in immunotherapy, especially in the treatment of arthritis
PT   -
XX
PS   Claim 12; Page 39; 52pp; English.
XX
CC   The sequence represents a modified peptide derived from Human cartilage
CC   gp-39 protein epitope (amino acids 263-275) for use in immunotherapy.
CC   The peptides are useful for inducing tolerance induction in patients
CC   suffering from autoimmune diseases such as rheumatoid arthritis.
CC   The peptides provide antigen-specific, non-toxic immunosuppression
CC   therapy as a very attractive alternative to non-specific
CC   immunosuppression.
XX
SQ   Sequence   13 AA;

  Query Match             100.0%;  Score 62;  DB 22;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 6e-05;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RSFTLASSETGVG 13
              |||||||||||||
Db          1 RSFTLASSETGVG 13


RESULT 7
AAU03059
ID   AAU03059 standard; peptide; 13 AA.
XX
AC   AAU03059;
XX
DT   12-SEP-2001  (first entry)
XX
DE   Human cartilage gp-39 autoimmune peptide #15.
XX
KW   Human; autoimmune disease; epitope; antigen; rheumatoid arthritis;
KW   immunosuppression.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "N-terminal (N-methyl-nicotinoyl)+"
XX
PN   WO200129081-A1.
XX
PD   26-APR-2001.
XX
PF   12-OCT-2000; 2000WO-EP10230.
XX
PR   18-OCT-1999;   99EP-0203427.
XX
PA   (ALKU ) AKZO NOBEL NV.
XX
PI   Van Staveren CJ,  Timmers CM,  Van Galen PJM,  Knegtel RMA,  Boots AMH;
PI   Miltenburg AMM;
XX
DR   WPI; 2001-367293/38.
XX
PT   Modified peptides derived from human cartilage gp-39 protein epitope
PT   are useful in immunotherapy, especially in the treatment of arthritis
PT   -
XX
PS   Claim 12; Page 39; 52pp; English.
XX
CC   The sequence represents a modified peptide derived from Human cartilage
CC   gp-39 protein epitope (amino acids 263-275) for use in immunotherapy.
CC   The peptides are useful for inducing tolerance induction in patients
CC   suffering from autoimmune diseases such as rheumatoid arthritis.
CC   The peptides provide antigen-specific, non-toxic immunosuppression
CC   therapy as a very attractive alternative to non-specific
CC   immunosuppression.
XX
SQ   Sequence   13 AA;

  Query Match             100.0%;  Score 62;  DB 22;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 6e-05;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RSFTLASSETGVG 13
              |||||||||||||
Db          1 RSFTLASSETGVG 13


RESULT 8
AAU03065
ID   AAU03065 standard; peptide; 13 AA.
XX
AC   AAU03065;
XX
DT   12-SEP-2001  (first entry)
XX
DE   Human cartilage gp-39 autoimmune peptide #21.
XX
KW   Human; autoimmune disease; epitope; antigen; rheumatoid arthritis;
KW   immunosuppression.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "N-terminal acetyl"
FT   Misc-difference 1
FT                   /note= "D-from residue"
XX
PN   WO200129081-A1.
```

```
XX
PD   26-APR-2001.
XX
PF   12-OCT-2000; 2000WO-EP10230.
XX
PR   18-OCT-1999;   99EP-0203427.
XX
PA   (ALKU ) AKZO NOBEL NV.
XX
PI   Van Staveren CJ,  Timmers CM,  Van Galen PJM,  Knegtel RMA,  Boots AMH;
PI   Miltenburg AMM;
XX
DR   WPI; 2001-367293/38.
XX
PT   Modified peptides derived from human cartilage gp-39 protein epitope
PT   are useful in immunotherapy, especially in the treatment of arthritis
PT   -
XX
PS   Disclosure; Page 34; 52pp; English.
XX
CC   The sequence represents a modified peptide derived from Human cartilage
CC   gp-39 protein epitope (amino acids 263-275) for use in immunotherapy.
CC   The peptides are useful for inducing tolerance induction in patients
CC   suffering from autoimmune diseases such as rheumatoid arthritis.
CC   The peptides provide antigen-specific, non-toxic immunosuppression
CC   therapy as a very attractive alternative to non-specific
CC   immunosuppression.
XX
SQ   Sequence   13 AA;

  Query Match             100.0%;  Score 62;  DB 22;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 6e-05;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RSFTLASSETGVG 13
              |||||||||||||
Db          1 RSFTLASSETGVG 13


RESULT 9
AAU03066
ID   AAU03066 standard; peptide; 13 AA.
XX
AC   AAU03066;
XX
DT   12-SEP-2001  (first entry)
XX
DE   Human cartilage gp-39 autoimmune peptide #22.
XX
KW   Human; autoimmune disease; epitope; antigen; rheumatoid arthritis;
KW   immunosuppression.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "N-terminal acetyl"
XX
PN   WO200129081-A1.
XX
PD   26-APR-2001.
XX
PF   12-OCT-2000; 2000WO-EP10230.
XX
PR   18-OCT-1999;   99EP-0203427.
XX
PA   (ALKU ) AKZO NOBEL NV.
XX
PI   Van Staveren CJ,  Timmers CM,  Van Galen PJM,  Knegtel RMA,  Boots AMH;
PI   Miltenburg AMM;
XX
DR   WPI; 2001-367293/38.
XX
PT   Modified peptides derived from human cartilage gp-39 protein epitope
PT   are useful in immunotherapy, especially in the treatment of arthritis
PT   -
XX
PS   Example 15; Page 32; 52pp; English.
XX
CC   The sequence represents a modified peptide derived from Human cartilage
CC   gp-39 protein epitope (amino acids 263-275) for use in immunotherapy.
CC   The peptides are useful for inducing tolerance induction in patients
CC   suffering from autoimmune diseases such as rheumatoid arthritis.
CC   The peptides provide antigen-specific, non-toxic immunosuppression
CC   therapy as a very attractive alternative to non-specific
CC   immunosuppression.
XX
SQ   Sequence   13 AA;

  Query Match             100.0%;  Score 62;  DB 22;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 6e-05;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RSFTLASSETGVG 13
              |||||||||||||
Db          1 RSFTLASSETGVG 13


RESULT 10
AAU03067
ID   AAU03067 standard; peptide; 13 AA.
XX
AC   AAU03067;
XX
DT   12-SEP-2001  (first entry)
XX
DE   Human cartilage gp-39 autoimmune peptide #23.
XX
KW   Human; autoimmune disease; epitope; antigen; rheumatoid arthritis;
KW   immunosuppression.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "N-terminal MeO-C(O)"
XX
PN   WO200129081-A1.
XX
PD   26-APR-2001.
XX
PF   12-OCT-2000; 2000WO-EP10230.
XX
PR   18-OCT-1999;   99EP-0203427.
XX
PA   (ALKU ) AKZO NOBEL NV.
XX
PI   Van Staveren CJ,  Timmers CM,  Van Galen PJM,  Knegtel RMA,  Boots AMH;
PI   Miltenburg AMM;
XX
DR   WPI; 2001-367293/38.
XX
PT   Modified peptides derived from human cartilage gp-39 protein epitope
PT   are useful in immunotherapy, especially in the treatment of arthritis
PT   -
XX
PS   Example 9; Page 25; 52pp; English.
XX
CC   The sequence represents a modified peptide derived from Human cartilage
CC   gp-39 protein epitope (amino acids 263-275) for use in immunotherapy.
CC   The peptides are useful for inducing tolerance induction in patients
CC   suffering from autoimmune diseases such as rheumatoid arthritis.
CC   The peptides provide antigen-specific, non-toxic immunosuppression
CC   therapy as a very attractive alternative to non-specific
```

```
CC   immunosuppression..
XX
SQ   Sequence   13 AA;

  Query Match           100.0%;  Score 62;  DB 22;  Length 13;
  Best Local Similarity  100.0%;  Pred. No. 6e-05;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RSFTLASSETGVG 13
              |||||||||||||
Db          1 RSFTLASSETGVG 13


RESULT 11
AAE20207
ID   AAE20207 standard; peptide; 13 AA.
XX
AC   AAE20207;
XX
DT   18-JUN-2002  (first entry)
XX
DE   Human cartilage glycoprotein (HC gp)-39 peptide (residues 263-275).
XX
KW   Human cartilage glycoprotein-39; HC gp-39; autoimmune disorder; MHC;
KW   major histocompatibility complex; insulin-dependent diabetes mellitus;
KW   multiple sclerosis; psoriasis; myasthenia gravis; rheumatoid arthritis;
KW   immunosuppressive; neuroprotective; therapy.
XX
OS   Homo sapiens.
XX
PN   WO200214870-A2.
XX
PD   21-FEB-2002.
XX
PF   08-AUG-2001; 2001WO-EP09136.
XX
PR   14-AUG-2000; 2000EP-0202844.
XX
PA   (ALKU ) AKZO NOBEL NV.
XX
PI   Steenbakkers PGA;
XX
DR   WPI; 2002-269214/31.
XX
PT   Method for diagnosing an autoimmune disease activity by detecting the
PT   presence of an autoimmune specific major histocompatibility complex
PT   (MHC)-peptide complex in a patient using antibodies specific against
PT   MHC-peptide complexes -
XX
PS   Example 4; Page 23; 43pp; English.
XX
CC   The patent discloses a method for diagnosing an autoimmune disease
CC   activity. The method involves detecting the presence of an autoimmune
CC   specific major histocompatibility complex (MHC)-peptide complex in a
CC   patient suffering from an autoimmune disease with antibodies or its
CC   antigen binding domains that specifically bind to a MHC-gp39-derived
CC   peptide. The method and antibodies are useful for diagnosing auto-
CC   immune diseases or purifying autoimmune specific MHC-peptide complexes.
CC   They are also useful in therapy or for the manufacture of pharmaceutical
CC   preparation for the treatment of autoimmune disorders such as insulin-
CC   dependent diabetes mellitus, multiple sclerosis, myasthenia gravis,
CC   psoriasis or rheumatoid arthritis. The present sequence is human
CC   cartilage glycoprotein (HC gp)-39 peptide (residues 263-275). This
CC   peptide is used in the exemplification of the invention.
XX
SQ   Sequence   13 AA;

  Query Match           100.0%;  Score 62;  DB 23;  Length 13;
  Best Local Similarity  100.0%;  Pred. No. 6e-05;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RSFTLASSETGVG 13
              |||||||||||||
Db          1 RSFTLASSETGVG 13


RESULT 12
AAE20214
ID   AAE20214 standard; peptide; 13 AA.
XX
AC   AAE20214;
XX
DT   18-JUN-2002  (first entry)
XX
DE   Human cartilage glycoprotein (HC gp)-39 modified peptide #1.
XX
KW   Human cartilage glycoprotein-39; HC gp-39; autoimmune disorder; MHC;
KW   major histocompatibility complex; insulin-dependent diabetes mellitus;
KW   multiple sclerosis; psoriasis; myasthenia gravis; rheumatoid arthritis;
KW   immunosuppressive; neuroprotective; therapy.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "N-terminal acetyl"
XX
PN   WO200214870-A2.
XX
PD   21-FEB-2002.
XX
PF   08-AUG-2001; 2001WO-EP09136.
XX
PR   14-AUG-2000; 2000EP-0202844.
XX
PA   (ALKU ) AKZO NOBEL NV.
XX
PI   Steenbakkers PGA;
XX
DR   WPI; 2002-269214/31.
XX
PT   Method for diagnosing an autoimmune disease activity by detecting the
PT   presence of an autoimmune specific major histocompatibility complex
PT   (MHC)-peptide complex in a patient using antibodies specific against
PT   MHC-peptide complexes -
XX
PS   Example 4; Page 23; 43pp; English.
XX
CC   The patent discloses a method for diagnosing an autoimmune disease
CC   activity. The method involves detecting the presence of an autoimmune
CC   specific major histocompatibility complex (MHC)-peptide complex in a
CC   patient suffering from an autoimmune disease with antibodies or its
CC   antigen binding domains that specifically bind to a MHC-gp39-derived
CC   peptide. The method and antibodies are useful for diagnosing auto-
CC   immune diseases or purifying autoimmune specific MHC-peptide complexes.
CC   They are also useful in therapy or for the manufacture of pharmaceutical
CC   preparation for the treatment of autoimmune disorders such as insulin-
CC   dependent diabetes mellitus, multiple sclerosis, myasthenia gravis,
CC   psoriasis or rheumatoid arthritis. The present sequence is human
CC   cartilage glycoprotein (HC gp)-39 modified peptide. This peptide is
CC   used in the exemplification of the invention.
XX
SQ   Sequence   13 AA;

  Query Match           100.0%;  Score 62;  DB 23;  Length 13;
  Best Local Similarity  100.0%;  Pred. No. 6e-05;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RSFTLASSETGVG 13
              |||||||||||||
Db          1 RSFTLASSETGVG 13


RESULT 13
AAE20215
ID   AAE20215 standard; peptide; 13 AA.
```

```
XX
AC   AAE20215;
XX
DT   18-JUN-2002  (first entry)
XX
DE   Human cartilage glycoprotein (HC gp)-39 modified peptide #2.
XX
KW   Human cartilage glycoprotein-39; HC gp-39; autoimmune disorder; MHC;
KW   major histocompatibility complex; insulin-dependent diabetes mellitus;
KW   multiple sclerosis; psoriasis; myasthenia gravis; rheumatoid arthritis;
KW   immunosuppressive; neuroprotective; therapy.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "Linked to HOCH2-(CHOH)4-(CH2)"
XX
PN   WO200214870-A2.
XX
PD   21-FEB-2002.
XX
PF   08-AUG-2001; 2001WO-EP09136.
XX
PR   14-AUG-2000; 2000EP-0202844.
XX
PA   (AKZO ) AKZO NOBEL NV.
XX
PI   Steenbakkers PGA;
XX
DR   WPI; 2002-269214/31.
XX
PT   Method for diagnosing an autoimmune disease activity by detecting the
PT   presence of an autoimmune specific major histocompatibility complex
PT   (MHC)-peptide complex in a patient using antibodies specific against
PT   MHC-peptide complexes -
XX
PS   Example 4; Page 23; 43pp; English.
XX
CC   The patent discloses a method for diagnosing an autoimmune disease
CC   activity. The method involves detecting the presence of an autoimmune
CC   specific major histocompatibility complex (MHC)-peptide complex in a
CC   patient suffering from an autoimmune disease with antibodies or its
CC   antigen binding domains that specifically bind to a MHC-gp39-derived
CC   peptide. The method and antibodies are useful for diagnosing auto-
CC   immune diseases or purifying autoimmune specific MHC-peptide complexes.
CC   They are also useful in therapy or for the manufacture of pharmaceutical
CC   preparation for the treatment of autoimmune disorders such as insulin-
CC   dependent diabetes mellitus, multiple sclerosis, myasthenia gravis,
CC   psoriasis or rheumatoid arthritis. The present sequence is human
CC   cartilage glycoprotein (HC gp)-39 modified peptide. This peptide is
CC   used in the exemplification of the invention.
XX
SQ   Sequence   13 AA;

  Query Match             100.0%;  Score 62;  DB 23;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 6e-05;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RSFTLASSETGVG 13
              |||||||||||||
Db          1 RSFTLASSETGVG 13


RESULT 14
AAE20216
ID   AAE20216 standard; peptide; 13 AA.
XX
AC   AAE20216;
XX
DT   18-JUN-2002  (first entry)
XX
DE   Human cartilage glycoprotein (HC gp)-39 modified peptide #3.
XX
KW   Human cartilage glycoprotein-39; HC gp-39; autoimmune disorder; MHC;
KW   major histocompatibility complex; insulin-dependent diabetes mellitus;
KW   multiple sclerosis; psoriasis; myasthenia gravis; rheumatoid arthritis;
KW   immunosuppressive; neuroprotective; therapy.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "N-terminal acetyl"
FT   Misc-difference 12
FT                   /note= "This residue is given as Xaa in the sequence
FT                   shown as SEQ ID NO: 11 in the sequence listing"
FT   Modified-site   12..13
FT                   /note= "Linked by psi-CH2NH group"
FT   Modified-site   13
FT                   /note= "C-terminal amide"
XX
PN   WO200214870-A2.
XX
PD   21-FEB-2002.
XX
PF   08-AUG-2001; 2001WO-EP09136.
XX
PR   14-AUG-2000; 2000EP-0202844.
XX
PA   (AKZO ) AKZO NOBEL NV.
XX
PI   Steenbakkers PGA;
XX
DR   WPI; 2002-269214/31.
XX
PT   Method for diagnosing an autoimmune disease activity by detecting the
PT   presence of an autoimmune specific major histocompatibility complex
PT   (MHC)-peptide complex in a patient using antibodies specific against
PT   MHC-peptide complexes -
XX
PS   Example 4; Page 23; 43pp; English.
XX
CC   The patent discloses a method for diagnosing an autoimmune disease
CC   activity. The method involves detecting the presence of an autoimmune
CC   specific major histocompatibility complex (MHC)-peptide complex in a
CC   patient suffering from an autoimmune disease with antibodies or its
CC   antigen binding domains that specifically bind to a MHC-gp39-derived
CC   peptide. The method and antibodies are useful for diagnosing auto-
CC   immune diseases or purifying autoimmune specific MHC-peptide complexes.
CC   They are also useful in therapy or for the manufacture of pharmaceutical
CC   preparation for the treatment of autoimmune disorders such as insulin-
CC   dependent diabetes mellitus, multiple sclerosis, myasthenia gravis,
CC   psoriasis or rheumatoid arthritis. The present sequence is human
CC   cartilage glycoprotein (HC gp)-39 modified peptide. This peptide is
CC   used in the exemplification of the invention.
XX
SQ   Sequence   13 AA;

  Query Match             100.0%;  Score 62;  DB 23;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 6e-05;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RSFTLASSETGVG 13
              |||||||||||||
Db          1 RSFTLASSETGVG 13


RESULT 15
AAE20220
ID   AAE20220 standard; peptide; 13 AA.
XX
AC   AAE20220;
XX
DT   18-JUN-2002  (first entry)
XX
```

```
DE   Human cartilage glycoprotein (HC gp)-39 modified peptide #6.
XX
KW   Human cartilage glycoprotein-39; HC gp-39; autoimmune disorder; MHC;
KW   major histocompatibility complex; insulin-dependent diabetes mellitus;
KW   multiple sclerosis; psoriasis; myasthenia gravis; rheumatoid arthritis;
KW   immunosuppressive; neuroprotective; therapy.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "N-terminal acetyl"
FT   Modified-site   13
FT                   /note= "C-terminal amide"
XX
PN   WO200214870-A2.
XX
PD   21-FEB-2002.
XX
PF   08-AUG-2001; 2001WO-EP09136.
XX
PR   14-AUG-2000; 2000EP-0202844.
XX
PA   (ALKU ) AKZO NOBEL NV.
XX
PI   Steenbakkers PGA;
XX
DR   WPI; 2002-269214/31.
XX
PT   Method for diagnosing an autoimmune disease activity by detecting the
PT   presence of an autoimmune specific major histocompatibility complex
PT   (MHC)-peptide complex in a patient using antibodies specific against
PT   MHC-peptide complexes -
XX
PS   Example 4; Page 23; 43pp; English.
XX
CC   The patent discloses a method for diagnosing an autoimmune disease
CC   activity. The method involves detecting the presence of an autoimmune
CC   specific major histocompatibility complex (MHC)-peptide complex in a
CC   patient suffering from an autoimmune disease with antibodies or its
CC   antigen binding domains that specifically bind to a MHC-gp39-derived
CC   peptide. The method and antibodies are useful for diagnosing auto-
CC   immune diseases or purifying autoimmune specific MHC-peptide complexes.
CC   They are also useful in therapy or for the manufacture of pharmaceutical
CC   preparation for the treatment of autoimmune disorders such as insulin-
CC   dependent diabetes mellitus, multiple sclerosis, myasthenia gravis,
CC   psoriasis or rheumatoid arthritis. The present sequence is human
CC   cartilage glycoprotein (HC gp)-39 modified peptide. This peptide is
CC   used in the exemplification of the invention.
XX
SQ   Sequence   13 AA;

  Query Match             100.0%;  Score 62;  DB 23;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 6e-05;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RSFTLASSETGVG 13
              |||||||||||||
Db          1 RSFTLASSETGVG 13


Search completed: September 24, 2003, 13:41:22
Job time : 52.2711 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 24, 2003, 13:35:23 ; Search time 15.7753 Seconds
(without alignments)
34.867 Million cell updates/sec

Title: US-09-744-282-6
Perfect score: 62
Sequence: 1 RSFTLASSETGVG 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	62	100.0	13	1	US-08-619-645-6	Sequence 6, Appli
2	62	100.0	13	2	US-08-634-493-6	Sequence 6, Appli
3	62	100.0	13	4	US-09-405-745-1	Sequence 1, Appli
4	62	100.0	15	4	US-09-405-745-2	Sequence 2, Appli
5	62	100.0	16	3	US-09-171-705-41	Sequence 41, Appl
6	62	100.0	16	4	US-09-405-745-3	Sequence 3, Appli
7	51	82.3	383	4	US-09-459-749D-17	Sequence 17, Appl
8	46	74.2	385	2	US-08-694-915-2	Sequence 2, Appli
9	46	74.2	416	2	US-08-694-915-4	Sequence 4, Appli
10	46	74.2	423	4	US-08-850-348A-2	Sequence 2, Appli
11	43	69.4	9	1	US-08-619-645-2	Sequence 2, Appli
12	43	69.4	9	2	US-08-634-493-2	Sequence 2, Appli
13	43	69.4	125	1	US-08-478-039-69	Sequence 69, Appl
14	43	69.4	125	1	US-08-476-349A-69	Sequence 69, Appl
15	42	67.7	13	4	US-09-459-749D-5	Sequence 5, Appli
16	41	66.1	13	1	US-08-619-645-5	Sequence 5, Appli
17	41	66.1	13	2	US-08-634-493-5	Sequence 5, Appli
18	41	66.1	16	3	US-09-171-705-40	Sequence 40, Appl
19	40.5	65.3	373	3	US-09-039-198A-14	Sequence 14, Appl
20	40.5	65.3	373	3	US-09-039-198A-15	Sequence 15, Appl
21	40.5	65.3	373	4	US-08-877-599-14	Sequence 14, Appl
22	40.5	65.3	373	4	US-08-877-599-15	Sequence 15, Appl
23	40.5	65.3	373	4	US-09-267-574-14	Sequence 14, Appl
24	40.5	65.3	373	4	US-09-267-574-15	Sequence 15, Appl
25	40.5	65.3	387	2	US-08-486-839-6	Sequence 6, Appli
26	40.5	65.3	387	3	US-09-151-011-6	Sequence 6, Appli
27	40.5	65.3	387	4	US-09-343-623-6	Sequence 6, Appli
28	40.5	65.3	466	2	US-08-486-839-4	Sequence 4, Appli
29	40.5	65.3	466	3	US-09-151-011-4	Sequence 4, Appli
30	40.5	65.3	466	3	US-09-039-198A-2	Sequence 2, Appli
31	40.5	65.3	466	3	US-09-039-198A-4	Sequence 4, Appli
32	40.5	65.3	466	4	US-09-343-623-4	Sequence 4, Appli
33	40.5	65.3	466	4	US-08-877-599-2	Sequence 2, Appli
34	40.5	65.3	466	4	US-08-877-599-4	Sequence 4, Appli
35	40.5	65.3	466	4	US-09-267-574-2	Sequence 2, Appli
36	40.5	65.3	466	4	US-09-267-574-4	Sequence 4, Appli
37	38	61.3	16	3	US-09-171-705-42	Sequence 42, Appl
38	37	59.7	200	4	US-09-562-737-107	Sequence 107, App
39	37	59.7	471	4	US-09-252-991A-19101	Sequence 19101, A
40	37	59.7	502	3	US-09-413-814-106	Sequence 106, App
41	36	58.1	82	2	US-08-773-251-22	Sequence 22, Appl
42	36	58.1	125	2	US-08-773-251-2	Sequence 2, Appli
43	35	56.5	439	4	US-08-311-731A-178	Sequence 178, App
44	35	56.5	481	4	US-08-976-063E-28	Sequence 28, Appl
45	35	56.5	685	2	US-08-849-480A-4	Sequence 4, Appli

ALIGNMENTS

```
RESULT 1
US-08-619-645-6
; Sequence 6, Application US/08619645
; Patent No. 5736507
;  GENERAL INFORMATION:
;    APPLICANT:  Boots, Anna M.H.
;    APPLICANT:  Verheijden, Gijsbertus F.M.
;    TITLE OF INVENTION:  No. 5736507el peptides derived from
;    TITLE OF INVENTION:  autoantigen for use in immunotherapy of autoimmune
;    TITLE OF INVENTION:  disease
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Akzo No. 5736507el Patent Department
;      STREET:  1300 Piccard Drive, Suite 206
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  U.S.A.
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25(EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/619,645
;      FILING DATE:  25-MAR-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Germley, Mary E.
;      REGISTRATION NUMBER:  34,409
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 947-4433
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  13 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-619-645-6

  Query Match           100.0%;  Score 62;  DB 1;  Length 13;
  Best Local Similarity  100.0%;  Pred. No. 1.3e-05;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RSFTLASSETGVG 13
              |||||||||||||
Db          1 RSFTLASSETGVG 13
```

```
RESULT 2
US-08-634-493-6
; Sequence 6, Application US/08634493
; Patent No. 5843449
;  GENERAL INFORMATION:
;    APPLICANT:  A.M.H. Boots
;    APPLICANT:  G.F.M. Verheijden
;    APPLICANT:  E.S. Bos
;    TITLE OF INVENTION:  No. 5843449el Peptides derived from autoantigen for use
;    TITLE OF INVENTION:  in Immunotherapy of Autoimmune Diseases
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Akzo No. 5843449el Patent Department
;      STREET:  1300 Piccard Drive, Suite 206
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/634,493
;      FILING DATE:  18-APR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/619,645
;      FILING DATE:  25-MAR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/EP95/04201
;      FILING DATE:  25-OCT-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  NL 9402031287
;      FILING DATE:  27-OCT-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  NL 952008860
;      FILING DATE:  07-APR-1995
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 258-5200
;      TELEFAX:  (301) 977-0847
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  13 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-634-493-6

  Query Match           100.0%;  Score 62;  DB 2;  Length 13;
  Best Local Similarity  100.0%;  Pred. No. 1.3e-05;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RSFTLASSETGVG 13
              |||||||||||||
Db          1 RSFTLASSETGVG 13


RESULT 3
US-09-405-745-1
; Sequence 1, Application US/09405745
; Patent No. 6392020
;  GENERAL INFORMATION:
;    APPLICANT:  Steenbakkers PGA
;    TITLE OF INVENTION:  Method of preparing a monoclonal
;    TITLE OF INVENTION:  antibody (as amended)
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Akzo No. 6392020el Patent Dept.
;      STREET:  1300 Piccard Drive, Suite 206
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  US
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/405,745
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/985,898
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sullivan, Michael G.
;      REGISTRATION NUMBER:  35,377
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  301-948-7400
;      TELEFAX:  301-948-9751
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  13 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FRAGMENT TYPE:  internal
US-09-405-745-1

  Query Match           100.0%;  Score 62;  DB 4;  Length 13;
  Best Local Similarity  100.0%;  Pred. No. 1.3e-05;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RSFTLASSETGVG 13
              |||||||||||||
Db          1 RSFTLASSETGVG 13


RESULT 4
US-09-405-745-2
; Sequence 2, Application US/09405745
; Patent No. 6392020
;  GENERAL INFORMATION:
;    APPLICANT:  Steenbakkers PGA
;    TITLE OF INVENTION:  Method of preparing a monoclonal
;    TITLE OF INVENTION:  antibody (as amended)
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Akzo No. 6392020el Patent Dept.
;      STREET:  1300 Piccard Drive, Suite 206
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  US
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/405,745
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/985,898
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sullivan, Michael G.
;      REGISTRATION NUMBER:  35,377
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  301-948-7400
;      TELEFAX:  301-948-9751
```

```
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  15 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FRAGMENT TYPE:  internal
US-09-405-745-2

  Query Match             100.0%;  Score 62;  DB 4;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-05;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RSFTLASSETGVG 13
              |||||||||||||
Db          3 RSFTLASSETGVG 15


RESULT 5
US-09-171-705-41
; Sequence 41, Application US/09171705
; Patent No. 6184204
; GENERAL INFORMATION:
;  APPLICANT: BOOTS, ANNA M.H.
;  APPLICANT: VERHEIJDEN, GILBERTUS F.M.
;  TITLE OF INVENTION: NOVEL PEPTIDES SUITABLE FOR USE IN ANTIGEN SPECIFIC
;  TITLE OF INVENTION: IMMUNOSUPPRESSIVE THERAPY
;  FILE REFERENCE: O/96198 US
;  CURRENT APPLICATION NUMBER: US/09/171,705
;  CURRENT FILING DATE: 1998-02-09
;  NUMBER OF SEQ ID NOS: 78
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 41
;   LENGTH: 16
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: DERIVED FROM
;   OTHER INFORMATION: SEQUENCE OF HUMAN CARTILAGE (HC) -39 PROTEIN
US-09-171-705-41

  Query Match             100.0%;  Score 62;  DB 3;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-05;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RSFTLASSETGVG 13
              |||||||||||||
Db          2 RSFTLASSETGVG 14


RESULT 6
US-09-405-745-3
; Sequence 3, Application US/09405745
; Patent No. 6392020
;  GENERAL INFORMATION:
;    APPLICANT:  Steenbakkers PGA
;    TITLE OF INVENTION:  Method of preparing a monoclonal
;    TITLE OF INVENTION:  antibody (as amended)
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Akzo No. 6392020el Patent Dept.
;      STREET:  1300 Piccard Drive, Suite 206
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  US
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/405,745
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/985,898
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sullivan, Michael G.
;      REGISTRATION NUMBER:  35,377
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  301-948-7400
;      TELEFAX:  301-948-9751
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  16 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FRAGMENT TYPE:  internal
US-09-405-745-3

  Query Match             100.0%;  Score 62;  DB 4;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-05;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RSFTLASSETGVG 13
              |||||||||||||
Db          2 RSFTLASSETGVG 14


RESULT 7
US-09-459-749D-17
; Sequence 17, Application US/09459749D
; Patent No. 6464975
; GENERAL INFORMATION:
;  APPLICANT: Millis, Albert J. T.
;  TITLE OF INVENTION: Compositions and Methods For Altering Cell Migration
;  FILE REFERENCE: 0794.016A
;  CURRENT APPLICATION NUMBER: US/09/459,749D
;  CURRENT FILING DATE:  1999-12-10
;  PRIOR APPLICATION NUMBER: 60/111,856
;  PRIOR FILING DATE: 1998-12-11
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
;   LENGTH: 383
;   TYPE: PRT
;   ORGANISM: Sus scrofa
US-09-459-749D-17

  Query Match             82.3%;  Score 51;  DB 4;  Length 383;
  Best Local Similarity   84.6%;  Pred. No. 0.084;
  Matches   11;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 RSFTLASSETGVG 13
              |||||||||:| ||
Db        263 RSFTLASSKIDVG 275


RESULT 8
US-08-694-915-2
; Sequence 2, Application US/08694915
; Patent No. 5811535
;  GENERAL INFORMATION:
;    APPLICANT:  Adamou, Julie
;    APPLICANT:  Kirkpatrick, Robert
;    APPLICANT:  Rosenberg, Martin
;    TITLE OF INVENTION:  HUMAN CARTILEGE GP39-LIKE GENE
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SmithKline Beecham Corporation
;      STREET:  709 Swedeland Road
```

```
;      CITY:  King of Prussia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19406-2799
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/694,915
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Han, William T
;      REGISTRATION NUMBER:  34,344
;      REFERENCE/DOCKET NUMBER:  ATG50017
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  610-270-5219
;      TELEFAX:  610-270-5090
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  385 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:  N-terminal
;    ORIGINAL SOURCE:
US-08-694-915-2

  Query Match           74.2%;  Score 46;  DB 2;  Length 385;
  Best Local Similarity  83.3%;  Pred. No. 0.75;
  Matches   10;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 SFTLASSETGVG 13
              |||:|||:|| ||
Db        267 SFTLASAETTVG 278


RESULT 9
US-08-694-915-4
; Sequence 4, Application US/08694915
; Patent No. 5811535
;  GENERAL INFORMATION:
;    APPLICANT:  Adamou, Julie
;    APPLICANT:  Kirkpatrick, Robert
;    APPLICANT:  Rosenberg, Martin
;    TITLE OF INVENTION:  HUMAN CARTILAGE GP39-LIKE GENE
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SmithKline Beecham Corporation
;      STREET:  709 Swedeland Road
;      CITY:  King of Prussia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19406-2799
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/694,915
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Han, William T
;      REGISTRATION NUMBER:  34,344
;      REFERENCE/DOCKET NUMBER:  ATG50017
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  610-270-5219
;      TELEFAX:  610-270-5090
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  416 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:  N-terminal
;    ORIGINAL SOURCE:
US-08-694-915-4

  Query Match           74.2%;  Score 46;  DB 2;  Length 416;
  Best Local Similarity  83.3%;  Pred. No. 0.83;
  Matches   10;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 SFTLASSETGVG 13
              |||:|||:|| ||
Db        298 SFTLASAETTVG 309


RESULT 10
US-08-850-348A-2
; Sequence 2, Application US/08850348A
; Patent No. 6576427
; GENERAL INFORMATION:
;  APPLICANT: Kirkpatrick, et al.
;  TITLE OF INVENTION: Human Cartilage Glycoprotein
;  FILE REFERENCE: PF498
;  CURRENT APPLICATION NUMBER: US/08/850,348A
;  CURRENT FILING DATE:  1997-05-02
;  PRIOR APPLICATION NUMBER: 60/016,532
;  PRIOR FILING DATE: 1996-05-03
;  NUMBER OF SEQ ID NOS: 3
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
;   LENGTH: 423
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-08-850-348A-2

  Query Match           74.2%;  Score 46;  DB 4;  Length 423;
  Best Local Similarity  83.3%;  Pred. No. 0.84;
  Matches   10;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 SFTLASSETGVG 13
              |||:|||:|| ||
Db        305 SFTLASAETTVG 316


RESULT 11
US-08-619-645-2
; Sequence 2, Application US/08619645
; Patent No. 5736507
;  GENERAL INFORMATION:
;    APPLICANT:  Boots, Anna M.H.
;    APPLICANT:  Verheijden, Gijsbertus F.M.
;    TITLE OF INVENTION:  No. 5736507el peptides derived from
;    TITLE OF INVENTION:  autoantigen for use in immunotherapy of autoimmune
;    TITLE OF INVENTION:  disease
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
```

```
;       ADDRESSEE:  Akzo No. 5736507el Patent Department
;       STREET:  1300 Piccard Drive, Suite 206
;       CITY:  Rockville
;       STATE:  Maryland
;       COUNTRY:  U.S.A.
;       ZIP:  20850
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25(EPO)
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/619,645
;       FILING DATE:  25-MAR-1996
;       CLASSIFICATION:  514
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Gormley, Mary E.
;       REGISTRATION NUMBER:  34,409
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (301) 947-4433
;   INFORMATION FOR SEQ ID NO:  2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  9 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-08-619-645-2

  Query Match             69.4%;  Score 43;  DB 1;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 FTLASSETG 11
              |||||||||
Db          1 FTLASSETG 9


RESULT 12
US-08-634-493-2
; Sequence 2, Application US/08634493
; Patent No. 5843449
;  GENERAL INFORMATION:
;    APPLICANT:  A.M.H. Boots
;    APPLICANT:  G.F.M. Verheijden
;    APPLICANT:  E.S. Bos
;    TITLE OF INVENTION:  No. 5843449el Peptides derived from autoantigen for use
;    TITLE OF INVENTION:  in Immunotherapy of Autoimmune Diseases
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Akzo No. 5843449el Patent Department
;      STREET:  1300 Piccard Drive, Suite 206
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/634,493
;      FILING DATE:  18-APR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/619,645
;      FILING DATE:  25-MAR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/EP95/04201
;      FILING DATE:  25-OCT-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  NL 942031287
;      FILING DATE:  27-OCT-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  NL 952006860
;      FILING DATE:  07-APR-1995
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 258-5200
;      TELEFAX:  (301) 977-0847
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-634-493-2

  Query Match             69.4%;  Score 43;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 FTLASSETG 11
              |||||||||
Db          1 FTLASSETG 9


RESULT 13
US-08-478-039-69
; Sequence 69, Application US/08478039
; Patent No. 5681722
;  GENERAL INFORMATION:
;    APPLICANT:  Newman, Roland A.
;    APPLICANT:  Hanna, Nabil
;    APPLICANT:  Raab, Ronald W.
;    TITLE OF INVENTION:  Recombinant Antibodies for Human Therapy
;    NUMBER OF SEQUENCES:  114
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BURNS, DOANE, SWECKER & MATHIS
;      STREET:  699 Prince St.
;      CITY:  Alexandria
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22313-1404
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/478,039
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/379,072
;      FILING DATE:  25-JAN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/912,292
;      FILING DATE:  10-JUL-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/856,281
;      FILING DATE:  23-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/735,064
;      FILING DATE:  25-JUL-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Teskin Esq., Robin L.
;      REGISTRATION NUMBER:  35,030
;      REFERENCE/DOCKET NUMBER:  012712-160
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6620
;      TELEFAX:  703-836-2021
;  INFORMATION FOR SEQ ID NO:  69:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  125 amino acids
;      TYPE:  amino acid
```

```
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  not relevant
;    MOLECULE TYPE:  peptide
;    ORIGINAL SOURCE:
;      ORGANISM:  Monkey
;    POSITION IN GENOME:
;      CHROMOSOME/SEGMENT:  VH2 clone 2-10
US-08-478-039-69

  Query Match             69.4%;  Score 43;  DB 1;  Length 125;
  Best Local Similarity   72.7%;  Pred. No. 0.75;
  Matches    8;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 FTLASSETGVG 13
              |:|:::|:|||
Db         27 FSLSTSETGVG 37


RESULT 14
US-08-476-349A-69
; Sequence 69, Application US/08476349A
; Patent No. 5750105
; GENERAL INFORMATION:
;    APPLICANT:  Newman, Roland A.
;    APPLICANT:  Hanna, Nabil
;    APPLICANT:  Raab, Ronald W.
;    TITLE OF INVENTION:  Recombinant Antibodies for Human Therapy
;    NUMBER OF SEQUENCES:  114
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BURNS, DOANE, SWECKER & MATHIS
;      STREET:  699 Prince St.
;      CITY:  Alexandria
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22313-1404
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/476,349A
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/379,272
;      FILING DATE:  25-JAN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/912,292
;      FILING DATE:  10-JUL-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/856,281
;      FILING DATE:  23-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/735,084
;      FILING DATE:  25-JUL-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Teskin Esq., Robin L.
;      REGISTRATION NUMBER:  35,030
;      REFERENCE/DOCKET NUMBER:  012712-161
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-836-6620
;      TELEFAX:  703-836-2021
;  INFORMATION FOR SEQ ID NO:  69:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  125 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  not relevant
;    MOLECULE TYPE:  peptide
;    ORIGINAL SOURCE:
;      ORGANISM:  Monkey
;    POSITION IN GENOME:
;      CHROMOSOME/SEGMENT:  VH2 clone 2-10
US-08-476-349A-69

  Query Match             69.4%;  Score 43;  DB 1;  Length 125;
  Best Local Similarity   72.7%;  Pred. No. 0.75;
  Matches    8;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 FTLASSETGVG 13
              |:|:::|:|||
Db         27 FSLSTSETGVG 37


RESULT 15
US-09-459-749D-5
; Sequence 5, Application US/09459749D
; Patent No. 6464975
; GENERAL INFORMATION:
;  APPLICANT: Millis, Albert J. T.
;  TITLE OF INVENTION: Compositions and Methods For Altering Cell Migration
;  FILE REFERENCE: 0794.015A
;  CURRENT APPLICATION NUMBER: US/09/459,749D
;  CURRENT FILING DATE:  1999-12-10
;  PRIOR APPLICATION NUMBER: 60/111,856
;  PRIOR FILING DATE: 1998-12-11
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
;  LENGTH: 13
;  TYPE: PRT
;  ORGANISM: Artificial Sequence
;  FEATURE:
;  OTHER INFORMATION: Description of Artificial Sequence:SITE
;  NAME/KEY: SITE
;  LOCATION: (1)..(13)
;  OTHER INFORMATION: Peptide corresponding to amino acids 261- 273 of
;  OTHER INFORMATION: gp38k protein
US-09-459-749D-5

  Query Match             67.7%;  Score 42;  DB 4;  Length 13;
  Best Local Similarity   90.0%;  Pred. No. 0.082;
  Matches    9;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RSFTLASSET 10
              |||||||||:|
Db          3 RSFTLASSKT 12


Search completed: September 24, 2003, 13:51:15
Job time : 15.7753 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 24, 2003, 13:47:40 ; Search time 26.0674 Seconds
(without alignments)
67.670 Million cell updates/sec

Title: US-09-744-282-6
Perfect score: 62
Sequence: 1 RSFTLASSETGVG 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 566894 seqs, 151307093 residues

Total number of hits satisfying chosen parameters: 566894

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	62	100.0	13	10	US-09-985-065-1	Sequence 1, Appli
2	62	100.0	15	10	US-09-985-065-2	Sequence 2, Appli
3	62	100.0	16	10	US-09-985-065-3	Sequence 3, Appli
4	62	100.0	199	9	US-09-815-837-56	Sequence 56, Appl
5	62	100.0	199	9	US-09-815-837-60	Sequence 60, Appl
6	62	100.0	208	9	US-09-815-837-102	Sequence 102, App
7	62	100.0	209	9	US-09-815-837-58	Sequence 58, Appl
8	62	100.0	210	9	US-09-815-837-40	Sequence 40, Appl
9	62	100.0	211	9	US-09-815-837-39	Sequence 39, Appl
10	62	100.0	211	9	US-09-815-837-59	Sequence 59, Appl
11	62	100.0	212	9	US-09-815-837-37	Sequence 37, Appl
12	62	100.0	212	9	US-09-815-837-38	Sequence 38, Appl
13	62	100.0	213	9	US-09-815-837-41	Sequence 41, Appl
14	62	100.0	213	9	US-09-815-837-43	Sequence 43, Appl
15	62	100.0	213	9	US-09-815-837-44	Sequence 44, Appl
16	62	100.0	213	9	US-09-815-837-57	Sequence 57, Appl
17	62	100.0	214	9	US-09-815-837-42	Sequence 42, Appl
18	62	100.0	216	9	US-09-815-837-45	Sequence 45, Appl
19	62	100.0	216	9	US-09-815-837-46	Sequence 46, Appl
20	62	100.0	220	9	US-09-815-837-47	Sequence 47, Appl
21	62	100.0	220	9	US-09-815-837-48	Sequence 48, Appl
22	62	100.0	383	15	US-10-097-340-45	Sequence 45, Appl
23	51	82.3	383	10	US-09-459-749D-17	Sequence 17, Appl
24	42	67.7	13	10	US-09-459-749D-5	Sequence 5, Appli
25	40.5	65.3	373	12	US-10-161-547-14	Sequence 14, Appl
26	40.5	65.3	373	12	US-10-161-547-15	Sequence 15, Appl
27	40.5	65.3	445	15	US-10-004-219B-10	Sequence 10, Appl
28	40.5	65.3	466	12	US-10-161-547-2	Sequence 2, Appli
29	40.5	65.3	466	12	US-10-161-547-4	Sequence 4, Appli
30	40	64.5	480	12	US-10-238-075-697	Sequence 697, App
31	39	62.9	184	9	US-09-867-550-938	Sequence 938, App
32	38	61.3	125	10	US-09-850-365-77	Sequence 77, Appl
33	38	61.3	228	11	US-09-972-656-84	Sequence 84, Appl
34	38	61.3	434	15	US-10-156-761-12513	Sequence 12513, A
35	37	59.7	200	15	US-10-211-962-107	Sequence 107, App
36	37	59.7	309	12	US-10-238-075-580	Sequence 580, App
37	37	59.7	890	15	US-10-171-311-196	Sequence 196, App
38	36	58.1	453	15	US-10-128-714-3161	Sequence 3161, Ap
39	36	58.1	453	15	US-10-128-714-8161	Sequence 8161, Ap
40	36	58.1	664	9	US-09-815-242-13820	Sequence 13820, A
41	36	58.1	4809	15	US-10-156-761-9090	Sequence 9090, Ap
42	35	56.5	199	15	US-10-156-761-8757	Sequence 8757, Ap
43	35	56.5	248	11	US-09-964-275B-6	Sequence 6, Appli
44	35	56.5	287	9	US-09-833-745-3	Sequence 3, Appli
45	35	56.5	296	11	US-09-964-275B-2	Sequence 2, Appli

ALIGNMENTS

```
RESULT 1
US-09-985-065-1
; Sequence 1, Application US/09985065
; Patent No. US20020143150A1
;  GENERAL INFORMATION:
;       APPLICANT: Steenbakkers PGA
;       TITLE OF INVENTION: Method of preparing a monoclonal
;                           antibody (as amended)
;       NUMBER OF SEQUENCES: 3
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Akzo No. US20020143150A1e1 Patent Dept.
;            STREET: 1300 Piccard Drive, Suite 206
;            CITY: Rockville
;            STATE: Maryland
;            COUNTRY: US
;            ZIP: 20850
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/985,065
;            FILING DATE: 01-No. US20020143150A1-2001
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 09/405,745
;            FILING DATE: 01-NOV-2001
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Sullivan, Michael G.
;            REGISTRATION NUMBER: 35,377
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 301-948-7400
;            TELEFAX: 301-948-9751
;  INFORMATION FOR SEQ ID NO: 1:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 13 amino acids
;            TYPE: amino acid
```

```
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;          MOLECULE TYPE: protein
;          FRAGMENT TYPE: internal
;          SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-985-065-1

  Query Match           100.0%;  Score 62;  DB 15;  Length 13;
  Best Local Similarity  100.0%;  Pred. No. 4e-05;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RSFTLASSEIGVG 13
              |||||||||||||
Db          1 RSFTLASSETGVG 13


RESULT 2
US-09-985-065-2
; Sequence 2, Application US/09985065
; Patent No. US20020143150A1
;   GENERAL INFORMATION:
;        APPLICANT: Steenbakkers PGA
;        TITLE OF INVENTION: Method of preparing a monoclonal
;                            antibody (as amended)
;        NUMBER OF SEQUENCES: 3
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Akzo No. US20020143150A1e1 Patent Dept.
;             STREET: 1300 Piccard Drive, Suite 206
;             CITY: Rockville
;             STATE: Maryland
;             COUNTRY: US
;             ZIP: 20850
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/985,065
;             FILING DATE: 01-No. US20020143150A1-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/405,745
;             FILING DATE: 01-NOV-2001
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Sullivan, Michael G.
;             REGISTRATION NUMBER: 35,377
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 301-948-7400
;             TELEFAX: 301-948-9751
;    INFORMATION FOR SEQ ID NO: 2:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 15 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        FRAGMENT TYPE: internal
;        SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-985-065-2

  Query Match           100.0%;  Score 62;  DB 13;  Length 15;
  Best Local Similarity  100.0%;  Pred. No. 4.7e-05;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RSFTLASSETGVG 13
              |||||||||||||
Db          3 RSFTLASSETGVG 15


RESULT 3
US-09-985-065-3
; Sequence 3, Application US/09985065
; Patent No. US20020143150A1
;   GENERAL INFORMATION:
;        APPLICANT: Steenbakkers PGA
;        TITLE OF INVENTION: Method of preparing a monoclonal
;                            antibody (as amended)
;        NUMBER OF SEQUENCES: 3
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Akzo No. US20020143150A1e1 Patent Dept.
;             STREET: 1300 Piccard Drive, Suite 206
;             CITY: Rockville
;             STATE: Maryland
;             COUNTRY: US
;             ZIP: 20850
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/985,065
;             FILING DATE: 01-No. US20020143150A1-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/405,745
;             FILING DATE: 01-NOV-2001
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Sullivan, Michael G.
;             REGISTRATION NUMBER: 35,377
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 301-948-7400
;             TELEFAX: 301-948-9751
;    INFORMATION FOR SEQ ID NO: 3:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 16 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        FRAGMENT TYPE: internal
;        SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-985-065-3

  Query Match           100.0%;  Score 62;  DB 13;  Length 16;
  Best Local Similarity  100.0%;  Pred. No. 5.1e-05;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RSFTLASSEIGVG 13
              |||||||||||||
Db          2 RSFTLASSETGVG 14


RESULT 4
US-09-815-837-56
; Sequence 56, Application US/09815837
; Patent No. US20020082411A1
; GENERAL INFORMATION:
;  APPLICANT: Carter, Darrick
;  APPLICANT:  Zhu, Shirley
;  APPLICANT:  Arimilli, Subhashini
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Immune Medators and Related Methods
;  FILE REFERENCE: 014058-005670US
;  CURRENT APPLICATION NUMBER: US/09/815,837
;  CURRENT FILING DATE:  2001-03-22
;  PRIOR APPLICATION NUMBER: US 60/191,274
;  PRIOR FILING DATE: 2000-03-22
;  PRIOR APPLICATION NUMBER: US 60/204,249
;  PRIOR FILING DATE: 2000-05-15
;  PRIOR APPLICATION NUMBER: US 60/264,003
;  PRIOR FILING DATE: 2001-01-23
;  NUMBER OF SEQ ID NOS: 129
;  SOFTWARE: PatentIn Ver. 2.1
```

; SEQ ID NO 56
; LENGTH: 199
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:CO528-AC single
; OTHER INFORMATION: chain recombinant MHC class II molecule
US-09-815-837-56

Query Match 100.0%; Score 62; DB 9; Length 199;
Best Local Similarity 100.0%; Pred. No. 0.0009;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RSFTLASSETGVG 13
||||||||||||
Db 6 RSFTLASSETGVG 18

RESULT 5
US-09-815-837-60
; Sequence 60, Application US/09815837
; Patent No. US20020082411A1
; GENERAL INFORMATION:
; APPLICANT: Carter, Darrick
; APPLICANT: Zhu, Shirley
; APPLICANT: Arimilli, Subhashini
; APPLICANT: Wang, Aijun
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Immune Medators and Related Methods
; FILE REFERENCE: 014058-005670US
; CURRENT APPLICATION NUMBER: US/09/815,837
; CURRENT FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: US 60/191,274
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: US 60/204,249
; PRIOR FILING DATE: 2000-05-15
; PRIOR APPLICATION NUMBER: US 60/264,003
; PRIOR FILING DATE: 2001-01-23
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 60
; LENGTH: 199
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:CO528 single
; OTHER INFORMATION: chain recombinant MHC class II molecule
US-09-815-837-60

Query Match 100.0%; Score 62; DB 9; Length 199;
Best Local Similarity 100.0%; Pred. No. 0.0009;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RSFTLASSETGVG 13
||||||||||||
Db 6 RSFTLASSETGVG 18

RESULT 6
US-09-815-837-102
; Sequence 102, Application US/09815837
; Patent No. US20020082411A1
; GENERAL INFORMATION:
; APPLICANT: Carter, Darrick
; APPLICANT: Zhu, Shirley
; APPLICANT: Arimilli, Subhashini
; APPLICANT: Wang, Aijun
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Immune Medators and Related Methods
; FILE REFERENCE: 014058-005670US
; CURRENT APPLICATION NUMBER: US/09/815,837
; CURRENT FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: US 60/191,274
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: US 60/204,249
; PRIOR FILING DATE: 2000-05-15
; PRIOR APPLICATION NUMBER: US 60/264,003
; PRIOR FILING DATE: 2001-01-23
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 102
; LENGTH: 208
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial
; OTHER INFORMATION: Sequence:qp39-beta1/alpha1 HLA-DR4 CO563 human
; OTHER INFORMATION: single chain molecule
US-09-815-837-102

Query Match 100.0%; Score 62; DB 9; Length 208;
Best Local Similarity 100.0%; Pred. No. 0.00095;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RSFTLASSETGVG 13
||||||||||||
Db 6 RSFTLASSETGVG 18

RESULT 7
US-09-815-837-58
; Sequence 58, Application US/09815837
; Patent No. US20020082411A1
; GENERAL INFORMATION:
; APPLICANT: Carter, Darrick
; APPLICANT: Zhu, Shirley
; APPLICANT: Arimilli, Subhashini
; APPLICANT: Wang, Aijun
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Immune Modators and Related Methods
; FILE REFERENCE: 014058-005670US
; CURRENT APPLICATION NUMBER: US/09/815,837
; CURRENT FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: US 60/191,274
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: US 60/204,249
; PRIOR FILING DATE: 2000-05-15
; PRIOR APPLICATION NUMBER: US 60/264,003
; PRIOR FILING DATE: 2001-01-23
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 58
; LENGTH: 209
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:CO698 single
; OTHER INFORMATION: chain recombinant MHC class II molecule
US-09-815-837-58

Query Match 100.0%; Score 62; DB 9; Length 209;
Best Local Similarity 100.0%; Pred. No. 0.00095;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RSFTLASSETGVG 13
||||||||||||
Db 5 RSFTLASSETGVG 17

RESULT 8
US-09-815-837-40
; Sequence 40, Application US/09815837
; Patent No. US20020082411A1
; GENERAL INFORMATION:
; APPLICANT: Carter, Darrick
; APPLICANT: Zhu, Shirley

; APPLICANT: Arimilli, Subhashini
; APPLICANT: Wang, Aijun
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Immune Medators and Related Methods
; FILE REFERENCE: 014058-005670US
; CURRENT APPLICATION NUMBER: US/09/815,837
; CURRENT FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: US 60/191,274
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: US 60/204,249
; PRIOR FILING DATE: 2000-05-15
; PRIOR APPLICATION NUMBER: US 60/264,003
; PRIOR FILING DATE: 2001-01-23
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 40
; LENGTH: 210
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:CO595 single
; OTHER INFORMATION: chain recombinant MHC class II molecule
US-09-815-837-40

Query Match 100.0%; Score 62; DB 9; Length 210;
Best Local Similarity 100.0%; Pred. No. 0.00096;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RSFTLASSETGVG 13
Db 6 RSFTLASSETGVG 18

RESULT 9
US-09-815-837-39
; Sequence 39, Application US/09815837
; Patent No. US20020082411A1
; GENERAL INFORMATION:
; APPLICANT: Carter, Darrick
; APPLICANT: Zhu, Shirley
; APPLICANT: Arimilli, Subhashini
; APPLICANT: Wang, Aijun
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Immune Medators and Related Methods
; FILE REFERENCE: 014058-005670US
; CURRENT APPLICATION NUMBER: US/09/815,837
; CURRENT FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: US 60/191,274
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: US 60/204,249
; PRIOR FILING DATE: 2000-05-15
; PRIOR APPLICATION NUMBER: US 60/264,003
; PRIOR FILING DATE: 2001-01-23
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 39
; LENGTH: 211
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:CO596 single
; OTHER INFORMATION: chain recombinant MHC class II molecule
US-09-815-837-39

Query Match 100.0%; Score 62; DB 9; Length 211;
Best Local Similarity 100.0%; Pred. No. 0.00097;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RSFTLASSETGVG 13
Db 6 RSFTLASSETGVG 18

RESULT 10
US-09-815-837-59
; Sequence 59, Application US/09815837
; Patent No. US20020082411A1
; GENERAL INFORMATION:
; APPLICANT: Carter, Darrick
; APPLICANT: Zhu, Shirley
; APPLICANT: Arimilli, Subhashini
; APPLICANT: Wang, Aijun
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Immune Medators and Related Methods
; FILE REFERENCE: 014058-005670US
; CURRENT APPLICATION NUMBER: US/09/815,837
; CURRENT FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: US 60/191,274
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: US 60/204,249
; PRIOR FILING DATE: 2000-05-15
; PRIOR APPLICATION NUMBER: US 60/264,003
; PRIOR FILING DATE: 2001-01-23
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 59
; LENGTH: 211
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:CO608 variation
; OTHER INFORMATION: single chain recombinant MHC class II molecule
US-09-815-837-59

Query Match 100.0%; Score 62; DB 9; Length 211;
Best Local Similarity 100.0%; Pred. No. 0.00097;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RSFTLASSETGVG 13
Db 6 RSFTLASSETGVG 18

RESULT 11
US-09-815-837-37
; Sequence 37, Application US/09815837
; Patent No. US20020082411A1
; GENERAL INFORMATION:
; APPLICANT: Carter, Darrick
; APPLICANT: Zhu, Shirley
; APPLICANT: Arimilli, Subhashini
; APPLICANT: Wang, Aijun
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Immune Medators and Related Methods
; FILE REFERENCE: 014058-005670US
; CURRENT APPLICATION NUMBER: US/09/815,837
; CURRENT FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: US 60/191,274
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: US 60/204,249
; PRIOR FILING DATE: 2000-05-15
; PRIOR APPLICATION NUMBER: US 60/264,003
; PRIOR FILING DATE: 2001-01-23
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 37
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:CO598 single
; OTHER INFORMATION: chain recombinant MHC class II molecule, human MHC
; OTHER INFORMATION: class II single chain molecule with murine linkers
US-09-815-837-37

Query Match 100.0%; Score 62; DB 9; Length 212;

```
  Best Local Similarity   100.0%;  Pred. No. 0.00097;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RSFTLASSETGVG 13
              |||||||||||||
Db          6 RSFTLASSETGVG 18


RESULT 12
US-09-815-837-38
; Sequence 38, Application US/09815837
; Patent No. US20020082411A1
; GENERAL INFORMATION:
;  APPLICANT: Carter, Darrick
;  APPLICANT:  Zhu, Shirley
;  APPLICANT:  Arimilli, Subhashini
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Immune Medators and Related Methods
;  FILE REFERENCE: 014058-005670US
;  CURRENT APPLICATION NUMBER: US/09/815,837
;  CURRENT FILING DATE:  2001-03-22
;  PRIOR APPLICATION NUMBER: US 60/191,274
;  PRIOR FILING DATE: 2000-03-22
;  PRIOR APPLICATION NUMBER: US 60/204,249
;  PRIOR FILING DATE: 2000-05-15
;  PRIOR APPLICATION NUMBER: US 60/264,003
;  PRIOR FILING DATE: 2001-01-23
;  NUMBER OF SEQ ID NOS: 129
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 38
;   LENGTH: 212
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:CO597 single
;   OTHER INFORMATION: chain recombinant MHC class II molecule
US-09-815-837-38

  Query Match             100.0%;  Score 62;  DB 9;  Length 212;
  Best Local Similarity   100.0%;  Pred. No. 0.00097;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RSFTLASSETGVG 13
              |||||||||||||
Db          6 RSFTLASSETGVG 18


RESULT 13
US-09-815-837-41
; Sequence 41, Application US/09815837
; Patent No. US20020082411A1
; GENERAL INFORMATION:
;  APPLICANT: Carter, Darrick
;  APPLICANT:  Zhu, Shirley
;  APPLICANT:  Arimilli, Subhashini
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Immune Medators and Related Methods
;  FILE REFERENCE: 014058-005670US
;  CURRENT APPLICATION NUMBER: US/09/815,837
;  CURRENT FILING DATE:  2001-03-22
;  PRIOR APPLICATION NUMBER: US 60/191,274
;  PRIOR FILING DATE: 2000-03-22
;  PRIOR APPLICATION NUMBER: US 60/204,249
;  PRIOR FILING DATE: 2000-05-15
;  PRIOR APPLICATION NUMBER: US 60/264,003
;  PRIOR FILING DATE: 2001-01-23
;  NUMBER OF SEQ ID NOS: 129
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41
;   LENGTH: 213
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:CO593 single
;   OTHER INFORMATION: chain recombinant MHC class II molecule
US-09-815-837-41

  Query Match             100.0%;  Score 62;  DB 9;  Length 213;
  Best Local Similarity   100.0%;  Pred. No. 0.00098;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RSFTLASSETGVG 13
              |||||||||||||
Db          6 RSFTLASSETGVG 18


RESULT 14
US-09-815-837-43
; Sequence 43, Application US/09815837
; Patent No. US20020082411A1
; GENERAL INFORMATION:
;  APPLICANT: Carter, Darrick
;  APPLICANT:  Zhu, Shirley
;  APPLICANT:  Arimilli, Subhashini
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Immune Medators and Related Methods
;  FILE REFERENCE: 014058-005670US
;  CURRENT APPLICATION NUMBER: US/09/815,837
;  CURRENT FILING DATE:  2001-03-22
;  PRIOR APPLICATION NUMBER: US 60/191,274
;  PRIOR FILING DATE: 2000-03-22
;  PRIOR APPLICATION NUMBER: US 60/204,249
;  PRIOR FILING DATE: 2000-05-15
;  PRIOR APPLICATION NUMBER: US 60/264,003
;  PRIOR FILING DATE: 2001-01-23
;  NUMBER OF SEQ ID NOS: 129
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 43
;   LENGTH: 213
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:CO594 single
;   OTHER INFORMATION: chain recombinant MHC class II molecule
US-09-815-837-43

  Query Match             100.0%;  Score 62;  DB 9;  Length 213;
  Best Local Similarity   100.0%;  Pred. No. 0.00098;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RSFTLASSETGVG 13
              |||||||||||||
Db          6 RSFTLASSETGVG 18


RESULT 15
US-09-815-837-44
; Sequence 44, Application US/09815837
; Patent No. US20020082411A1
; GENERAL INFORMATION:
;  APPLICANT: Carter, Darrick
;  APPLICANT:  Zhu, Shirley
;  APPLICANT:  Arimilli, Subhashini
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Immune Medators and Related Methods
;  FILE REFERENCE: 014058-005670US
;  CURRENT APPLICATION NUMBER: US/09/815,837
;  CURRENT FILING DATE:  2001-03-22
;  PRIOR APPLICATION NUMBER: US 60/191,274
;  PRIOR FILING DATE: 2000-03-22
;  PRIOR APPLICATION NUMBER: US 60/204,249
;  PRIOR FILING DATE: 2000-05-15
```

```
;  PRIOR APPLICATION NUMBER: US 60/264,003
;  PRIOR FILING DATE: 2001-01-23
;  NUMBER OF SEQ ID NOS: 129
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 44
;   LENGTH: 213
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:CO583 single
;   OTHER INFORMATION: chain recomoinant MHC class II molecule
US-09-815-837-44

  Query Match           100.0%;  Score 62;  DB 3;  Length 213;
  Best Local Similarity  100.0%;  Pred. No. 0.00098;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RSFTLASSEIGVG 13
              |||||||||||||
Db          6 RSFTLASSETGVG 18


Search completed: September 24, 2003, 14:22:29
Job time : 29.0674 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 24, 2003, 13:34:08 ; Search time 15.7753 Seconds
(without alignments)
79.250 Million cell updates/sec

Title: US-09-744-282-6
Perfect score: 62
Sequence: 1 RSFTLASSETGVG 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	62	100.0	383	2	A49562	cartilage glycopro
2	51	82.3	383	2	S51327	heparin-binding gl
3	43	69.4	405	2	S61551	breast-regressing
4	41	66.1	830	2	I06578	1,4-alpha-glucan b
5	40	64.5	185	2	S50094	endodeoxyribonucle
6	40	64.5	461	2	G90791	hypothetical prote
7	40	64.5	467	2	D85652	hypothetical prote
8	39	62.9	587	2	G70502	probable recN prot
9	38	61.3	261	2	F84325	hypothetical prote
10	38	61.3	286	2	AC3035	transcription regu
11	38	61.3	286	2	H98250	lysR-type transcri
12	38	61.3	441	2	T31090	probable ATP-bindi
13	37	59.7	242	2	A69026	conserved hypothet
14	37	59.7	298	2	D95901	probable transcrip
15	37	59.7	459	2	F83895	metal-tetracycline
16	36	58.1	251	2	C69046	conserved hypothet
17	36	58.1	266	2	C70788	hypothetical prote
18	36	58.1	424	2	G81350	probable aminotran
19	36	58.1	664	2	AG1026	topoisomerase B [i
20	36	58.1	860	2	T49910	hypothetical prote
21	36	58.1	861	2	T49909	hypothetical prote
22	36	58.1	1555	2	T18688	hypothetical prote
23	36	58.1	1973	2	G89608	protein B0272.5 [i
24	36	58.1	1973	2	T28586	hypothetical prote
25	36	58.1	2305	2	T15577	hypothetical prote
26	36	58.1	2305	2	B89608	protein C23F12.1 [
27	35	56.5	187	2	S69542	hypothetical prote
28	35	56.5	271	2	H97230	alpha/beta superfa
29	35	56.5	299	2	AD3156	transcription regu
30	35	56.5	299	2	E98131	transcription regu
31	35	56.5	302	2	F95387	protein [imported
32	35	56.5	320	2	T24905	hypothetical prote
33	35	56.5	355	2	H97715	alanine racemase (
34	35	56.5	363	2	B97044	MRP-family ATP-bin
35	35	56.5	414	2	T43387	hypothetical prote
36	35	56.5	433	2	G84612	probable GTP cyclo
37	35	56.5	435	2	T45199	probable mrp prote
38	35	56.5	445	2	B40590	flagellar distal c
39	35	56.5	573	2	B71324	probable DNA repai
40	35	56.5	596	2	AC1328	two-component sens
41	35	56.5	664	2	AG2960	glycosidase Atu328
42	35	56.5	664	2	G98322	hypothetical prote
43	35	56.5	686	2	G64618	cadmium-transporti
44	35	56.5	686	2	B71895	probable heavy-met
45	34	54.8	194	2	A84031	GTP-binding protei

ALIGNMENTS

RESULT 1
A49562
cartilage glycoprotein gp39 precursor - human
N;Alternate names: 39K synovial protein
C;Species: Homo sapiens (man)
C;Date: 23-Mar-1995 #sequence_revision 23-Mar-1995 #text_change 22-Jun-1999
C;Accession: A49562; S10677; A33152
R;Hakala, B.E.; White, C.; Recklies, A.D.
J. Biol. Chem. 268, 25803-25810, 1993
A;Title: Human cartilage gp-39, a major secretory product of articular chondrocytes
A;Reference number: A49562; MUID:94064658; PMID:8245017
A;Accession: A49562
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-383 <HAK>
A;Cross-references: GB:M80927; NID:g348911; PIDN:AAA16074.1; PID:g348912
R;Nyirkos, P.; Golds, E.E.
Biochem. J. 269, 265-268, 1990
A;Title: Human synovial cells secrete a 39 kDa protein similar to a bovine mammary
A;Reference number: S10677; MUID:90328983; PMID:2375755
A;Accession: S10677
A;Molecule type: protein
A;Residues: 22-40,'X',42-45 <NY2>
C;Superfamily: Streptomyces chitinase ch140
C;Keywords: cartilage; extracellular protein; glycoprotein
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-383/Product: cartilage glycoprotein gp39 #status predicted <MAT>

Query Match 100.0%; Score 62; DB 2; Length 383;
Best Local Similarity 100.0%; Pred. No. 0.00035;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 RSFTLASSETGVG 13
              |||||||||||||
Db        263 RSFTLASSETGVG 275
```

RESULT 2
S51327
heparin-binding glycoprotein 38K - pig
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 22-Jun-1999
C;Accession: S51327
R;Shackelton, L.M.; Mann, D.M.; Millis, A.J.T.
submitted to the EMBL Data Library, January 1995
A;Description: Identification of a 38kDa heparin-binding glycoprotein (gp38k) in di
A;Reference number: S51327
A;Accession: S51327
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-383 <SHA>
A;Cross-references: EMBL:Z47803; NID:g634097; PIDN:CAA87764.1; PID:g634098

```
C;Superfamily: Streptomyces chitinase chi40

  Query Match            82.3%;  Score 51;  DB 2;  Length 383;
  Best Local Similarity   84.6%;  Pred. No. 6.05;
  Matches   11;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 RSFTLASSETGVG 13
              ||||:|||:|:| ||
Db        263 RSFTLASSKTDVG 275


RESULT 3
S61551
breast-regressing protein brp39 precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
C;Accession: S61551; S61550; I48271
R;Morrison, B.W.; Loder, P.
Oncogene 9, 3417-3426, 1994
A;Title: neu and ras initiate murine mammary tumors that share genetic markers generally
A;Reference number: I48271; MUID:95060797; PMID:7970706
A;Accession: S61551
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-405 <MOR1>
A;Cross-references: EMBL:X93035; NID:g1085065; PIDN:CAA63603.1; PID:g1085066
R;Morrison, B.W.
submitted to the EMBL Data Library, November 1995
A;Reference number: S61550
A;Accession: S61550
A;Molecule type: mRNA
A;Residues: 1-245,'I',247-330,'H',332-350,'MVWALDLDDFQGTCQPKEPFPLTNAIKDALA' <MOR2>
A;Cross-references: EMBL:X93035; NID:g1085065; PIDN:CAA63603.1; PID:g1085066
A;Note: the differences at the carboxyl end are due to a frameshift error
C;Genetics:
A;Gene: brp39
C;Superfamily: Streptomyces chitinase chi40
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-405/Product: breast-regressing protein brp39 #status predicted <MAT>

  Query Match            69.4%;  Score 43;  DB 2;  Length 405;
  Best Local Similarity   69.2%;  Pred. No. 2.1;
  Matches    9;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 RSFTLASSETGVG 13
              :|||||||| :|
Db        264 KSFTLASSENQLG 276


RESULT 4
T06578
1,4-alpha-glucan branching enzyme (EC 2.4.1.18) sbe1 precursor - wheat
N;Alternate names: starch branching enzyme 1
C;Species: Triticum aestivum (common wheat)
C;Date: 23-Apr-1999 #sequence_revision 25-Apr-1999 #text_change 20-Jun-2000
C;Accession: T06578
R;Chibbar, R.N.
submitted to the EMBL Data Library, April 1997
A;Reference number: Z15772
A;Accession: T06578
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-830 <CHI>
A;Cross-references: EMBL:Y12320; PIDN:CAA72987.1
A;Experimental source: cv. Fielder; kernels at 12 days post-anthesis
C;Genetics:
A;Gene: Sbe1
C;Function:
A;Description: catalyzes the transglycosylation of a terminal segment of a 1,4-alpha-D-g
A;Pathway: glycogen/starch biosynthesis
C;Superfamily: 1,4-alpha-glucan branching enzyme
C;Keywords: amyloplast; glycogen/starch biosynthesis; glycosyltransferase; hexosyltrans
F;1-61/Domain: transit peptide (amyloplast) #status predicted <TNP>
F;62-830/Product: 1,4-alpha-glucan branching enzyme sbe1 #status predicted <MAT>

  Query Match            66.1%;  Score 41;  DB 2;  Length 830;
  Best Local Similarity   53.8%;  Pred. No. 11;
  Matches    7;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 RSFTLASSETGVG 13
              | :|:|:::| |||
Db         66 RDYTMATAEDGVG 78


RESULT 5
S50094
endodeoxyribonuclease (EC 3.1.21.-) - phage SP82
C;Species: phage SP82
A;Note: host Bacillus subtilis
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 29-Sep-1999
C;Accession: S50094
R;Goodrich-Blair, H.; Shub, D.A.
Nucleic Acids Res. 22, 3715-3721, 1994
A;Title: The DNA polymerase genes of several HMU-bacteriophages have similar group I
A;Reference number: S50091; MUID:95023117; PMID:7937082
A;Accession: S50094
A;Molecule type: DNA
A;Residues: 1-185 <GOO>
A;Cross-references: EMBL:U04812; NID:g495451; PIDN:AAA56884.1; PID:g495453
C;Superfamily: endodeoxyribonuclease
C;Keywords: hydrolase

  Query Match            64.5%;  Score 40;  DB 2;  Length 185;
  Best Local Similarity   75.0%;  Pred. No. 3.4;
  Matches    9;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 RSFTLASSETGV 12
              || | || ||||
Db        144 RSITQASKETGV 155


RESULT 6
G90791
hypothetical protein ECs1303 [imported] - Escherichia coli (strain O157:H7, substrai
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
C;Accession: G90791
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han,
qasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: G90791
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-461 <HAY>
A;Cross-references: GB:BA000007; PIDN:BAB34726.1; PID:g13360763; GSPDB:GN00154
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECs1303
C;Superfamily: Escherichia coli hypothetical protein b2625

  Query Match            64.5%;  Score 40;  DB 2;  Length 461;
  Best Local Similarity   75.0%;  Pred. No. 9.2;
  Matches    9;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 RSFTLASSETGV 12
              ||| | ||| ||
Db        178 RSFALVSSEGGV 189


RESULT 7
D85652
hypothetical protein Z1565 [imported] - Escherichia coli (strain O157:H7, substrain
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
```

```
C;Accession: D85652; C85601
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: D85652
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-467 <STO>
A;Cross-references: GB:AE005174; NID:g12514439; PIDN:AAG55680.1; GSPDB:GN00145; UWGP:Z15
A;Experimental source: strain O157:H7, substrain EDL933
A;Accession: C85601
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-467 <ST2>
A;Cross-references: GB:AE005174; NID:g12513922; PIDN:AAG55271.1; GSPDB:GN00145; UWGP:Z11
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: Z1565; Z1126
C;Superfamily: Escherichia coli hypothetical protein b2625

  Query Match             64.5%;  Score 40;  DB 2;  Length 467;
  Best Local Similarity   75.0%;  Pred. No. 9.4;
  Matches    9;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 RSFTLASSETGV 12
              ||| : ||| .|
Db        184 RSFALVSSEGGV 195


RESULT 8
G70502
probable recN protein - Mycobacterium tuberculosis (strain H37Rv)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2003
C;Accession: G70502
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: G70502
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-587 <COL>
A;Cross-references: GB:Z98268; GB:AL123456; NID:g3261839; PIDN:CAB10953.1; PID:g2326739
A;Experimental source: strain H37Rv
C;Genetics:
A;Gene: recN
C;Superfamily: recN protein

  Query Match             62.9%;  Score 39;  DB 2;  Length 587;
  Best Local Similarity   53.8%;  Pred. No. 19;
  Matches    7;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 RSFTLASSETGVG 13
              | ||: : |:. :
Db         22 RGFTVLTGETGTG 34


RESULT 9
F84325
hypothetical protein Vng1735c [imported] - Halobacterium sp. NRC-1
C;Species: Halobacterium sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C;Accession: F84325
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo
 Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M
A;Title: Genome sequence of Halobacterium species NRC-1.
A;Reference number: A84160; MUID:20504483; PMID:11016950
A;Accession: F84325
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-261 <STO>
A;Cross-references: GB:AE004437; NID:g10581197; PIDN:AAG19970.1; GSPDB:GN00138
C;Genetics:
A;Gene: VNG1735C

  Query Match             61.3%;  Score 38;  DB 2;  Length 261;
  Best Local Similarity   63.6%;  Pred. No. 12;
  Matches    7;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 RSFTLASSETG 11
              |:|:|  |:|
Db         35 RIFTLVDSDTG 45


RESULT 10
AC3035
transcription regulator, LysR family Atu3889 [imported] - Agrobacterium tumefaciens
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
C;Accession: AC3035
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; W
eraqe, G.; Gillet, W.; Grant, C.; Guenthner, D.; Kutyavin, T.; Levy, R.; Li, M.; Mc
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-K
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; MUID:21608550; PMID:11743193
A;Accession: AC3035
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-286 <KUR>
A;Cross-references: GB:AE008689; PIDN:AAL44697.1; PID:g17742326; GSPDB:GN00187
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: Atu3889
A;Map position: linear chromosome
C;Superfamily: hypothetical protein b1675

  Query Match             61.3%;  Score 38;  DB 2;  Length 286;
  Best Local Similarity   58.3%;  Pred. No. 14;
  Matches    7;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 SFTLASSETGVG 13
              |||||:  |:|
Db         18 SFTLAAERLGIG 29


RESULT 11
H98250
lysR-type transcription regulator [imported] - Agrobacterium tumefaciens (strain C5
C;Species: Agrobacterium tumefaciens
C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 18-Nov-2002
C;Accession: H98250
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Gol
 A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacteriu
A;Reference number: A97359; MUID:21608551; PMID:11743194
A;Accession: H98250
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-286 <KUR>
A;Cross-references: GB:AE007870; PIDN:AAK89530.1; PID:g15159409; GSPDB:GN00170
C;Genetics:
A;Gene: AGR_L_1915
A;Map position: linear chromosome
```

```
C;Superfamily: hypothetical protein b1875

  Query Match             61.3%;  Score 38;  DB 2;  Length 286;
  Best Local Similarity   58.3%;  Pred. No. 14;
  Matches    7;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 SFTLASSETGVG 13
              :|:|||:   |:|
Db         18 SFTLAAERLGIG 29


RESULT 12
T31090
probable ATP-binding protein - Serratia marcescens
C;Species: Serratia marcescens
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 17-Mar-2000
C;Accession: T31090
R;Saigi, F.; Climent, N.; Pique, N.; Sanchez, C.; Merino, S.; Rubires, X.; Aguilar, A.;
J. Bacteriol. 181, 1883-1891, 1999
A;Title: Genetic analysis of the Serratia marcescens N28b O4 antigen gene cluster.
A;Reference number: Z20974; MUID:99173913; PMID:10074083
A;Accession: T31090
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-441 <SAI>
A;Cross-references: EMBL:AF036816; NID:g2828669; PID:g2828672; PIDN:AAC00182.1
C;Genetics:
A;Gene: wzt
C;Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology

  Query Match             61.3%;  Score 38;  DB 2;  Length 441;
  Best Local Similarity   88.9%;  Pred. No. 22;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 SFTLASSET 10
              :|||||:|:|
Db        331 NFTLASSET 339


RESULT 13
A69026
conserved hypothetical protein MTH1192 - Methanobacterium thermoautotrophicum (strain De
C;Species: Methanobacterium thermoautotrophicum
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
C;Accession: A69026
R;Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.
ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A;Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct
A;Reference number: A69000; MUID:98037514; PMID:9371463
A;Accession: A69026
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-242 <MTH>
A;Cross-references: GB:AE000887; GB:AE000666; NID:g2622289; PIDN:AAB85681.1; PID:g262230
A;Experimental source: strain Delta H
C;Genetics:
A;Gene: MTH1192
A;Start codon: GTG

  Query Match             59.7%;  Score 37;  DB 2;  Length 242;
  Best Local Similarity   63.6%;  Pred. No. 18;
  Matches    7;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 SFTLASSETGV 12
              :|:: | :|||
Db        113 SFTMVSPDTGV 123


RESULT 14
D95901
probable transcription regulator protein [imported] - Sinorhizobium meliloti (strain 102
C;Species: Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
C;Accession: D95901
R;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hc
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A;Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing
A;Reference number: A95842; MUID:21396508; PMID:11481431
A;Accession: D95901
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-298 <KUR>
A;Cross-references: GB:AL591985; PIDN:CAC48876.1; PID:g15140349; GSPDB:GN00167
A;Experimental source: strain 1021, megaplasmid pSymB
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hub
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kann, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lela
hebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: SMb20494
A;Genome: plasmid

  Query Match             59.7%;  Score 37;  DB 2;  Length 298;
  Best Local Similarity   58.3%;  Pred. No. 22;
  Matches    7;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 RSFTLASSETGV 12
              |:|||||: : |:
Db         17 RSFTLAAEDLGL 28


RESULT 15
F83895
metal-tetracycline/H+ antiporter tetB [imported] - Bacillus halodurans (strain C-125
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 06-Jan-2003
C;Accession: F83895
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.;
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: F83895
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-459 <STO>
A;Cross-references: GB:AP001513; GB:BA000004; NID:g10174345; PIDN:BAB05685.1; GSPDB:
A;Experimental source: strain C-125
C;Genetics:
A;Gene: tetB
C;Superfamily: multidrug-efflux transporter

  Query Match             59.7%;  Score 37;  DB 2;  Length 459;
  Best Local Similarity   66.7%;  Pred. No. 36;
  Matches    8;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 SFTLASSETGVG 13
              | |||  | |||
Db        371 SMILAKKEEGVG 382


Search completed: September 24, 2003, 13:49:24
Job time : 16.7753 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 24, 2003, 13:26:37 ; Search time 8.17378 Seconds
(without alignments)
74.739 Million cell updates/sec

Title: US-09-744-282-6
Perfect score: 62
Sequence: 1 RSFTLASSETGVG 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	62	100.0	383	1	C3L1_HUMAN	P36222 homo sapien
2	46	74.2	390	1	C3L2_HUMAN	Q15782 homo sapien
3	43	69.4	381	1	C3L1_MOUSE	Q01362 mus musculu
4	39	62.9	587	1	RECN_MYCTO	Q33197 mycobacteri
5	38	61.3	887	1	K20A_MOUSE	P97329 mus musculu
6	37	59.7	890	1	K20A_HUMAN	O95235 homo sapien
7	36	58.1	650	1	GCP4_DROME	Q9vku7 drosophila
8	35	56.5	187	1	YO34_BPHP1	P51738 bacteriopha
9	35	56.5	336	1	TRPD_RHILO	Q98me4 rhizobium l
10	35	56.5	355	1	ALR_RICCN	Q92j49 rickettsia
11	35	56.5	383	1	MSP_MYCLE	P53382 mycobacteri
12	35	56.5	402	1	CN7B_SCHPO	Q09722 schizosacch
13	35	56.5	445	1	FLDL_VIBPA	Q93475 vibrio para
14	35	56.5	573	1	RECN_TREPA	O83456 treponema p
15	35	56.5	686	1	HMCT_HELPJ	Q9z153 helicobacte
16	35	56.5	686	1	HMCT_HELPY	Q59465 helicobacte
17	34	54.8	194	1	ENGB_BACHD	Q9k8f7 bacillus ha
18	34	54.8	246	1	POL1_CHICK	P10399 gallus gall
19	34	54.8	340	1	TRPD_AQUAE	O66576 aquifex aeo
20	34	54.8	459	1	TRME_BUCBP	P59569 buchnera ap
21	34	54.8	539	1	FUT2_ARATH	O81053 arabidopsis
22	34	54.8	647	1	YIK1_YEAST	P40489 saccharomyc
23	34	54.8	782	1	SM4B_MOUSE	Q62179 mus musculu
24	34	54.8	3591	1	FHAB_BORPE	P12255 bordetella
25	33	53.2	96	1	TRAA_ECOLI	P27187 escherichia
26	33	53.2	204	1	YPB1_BACAN	P13422 bacillus an
27	33	53.2	360	1	YACK_BACSU	P37573 bacillus su
28	33	53.2	366	1	YCGT_ECOLI	P76015 escherichia
29	33	53.2	372	1	GPR_BACHD	Q9kd78 bacillus ha
30	33	53.2	449	1	TBA8_HUMAN	Q9ny65 homo sapien
31	33	53.2	449	1	TBA8_MOUSE	Q9jjz2 mus musculu
32	33	53.2	467	1	EUTE_ECOLI	P77445 escherichia
33	33	53.2	467	1	EUTE_SALTY	P41793 salmonella
34	33	53.2	572	1	RECN_STRCO	Q9s220 streptomyce
35	33	53.2	576	1	RECN_BACSU	P17894 bacillus su
36	33	53.2	587	1	RECN_MYCLE	Q49896 mycobacteri
37	33	53.2	625	1	HTPG_BACHD	Q9ke51 bacillus ha
38	33	53.2	626	1	HTPG_BACSU	P46208 bacillus su
39	33	53.2	1076	1	NOP1_YEAST	P20676 saccharomyc
40	32	51.6	104	1	YBK0_YEAST	P38168 saccharomyc
41	32	51.6	123	1	RL19_TREPA	O83879 treponema p
42	32	51.6	137	1	YUXK_BACSU	P40761 bacillus su
43	32	51.6	156	1	RISB_NEIMA	Q9jqy6 neisseria m
44	32	51.6	212	1	BIOD_AGRT5	Q8uBL9 agrobacteri
45	32	51.6	218	1	Y5A4_STRCO	O87km3 streptomyce

ALIGNMENTS

```
RESULT 1
C3L1_HUMAN
ID   C3L1_HUMAN     STANDARD;      PRT:   383 AA.
AC   P36222; P30923;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Chitinase-3 like protein 1 precursor (Cartilage glycoprotein-39) (GP-
DE   39) (39 kDa synovial protein) (YKL-40).
GN   CHI3L1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC   TISSUE=Cartilage;
RX   MEDLINE=94064658; PubMed=8245017;
RA   Hakala B.E., White C., Recklies A.D.;
RT   "Human cartilage gp-39, a major secretory product of articular
RT   chondrocytes and synovial cells, is a mammalian member of a chitinase
RT   protein family.";
RL   J. Biol. Chem. 268:25803-25810(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Blood;
RX   MEDLINE=97386591; PubMed=9244440;
RA   Rehli M., Krause S.W., Andreesen R.;
RT   "Molecular characterization of the gene for human cartilage gp-39
RT   (CHI3L1), a member of the chitinase protein family and marker for
RT   late stages of macrophage differentiation.";
RL   Genomics 43:221-225(1997).
RN   [3]
RP   SEQUENCE OF 22-45.
RX   MEDLINE=90328983; PubMed=2375755;
RA   Nyirkos P., Golds E.E.;
RT   "Human synovial cells secrete a 39 kDa protein similar to a bovine
RT   mammary protein expressed during the non-lactating period.";
RL   Biochem. J. 269:265-268(1990).
CC   -!- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN THE CAPACITY OF CELLS TO
CC       RESPOND TO AND COPE WITH CHANGES IN THEIR ENVIRONMENT.
CC   -!- SUBCELLULAR LOCATION: Extracellular.
CC   -!- TISSUE SPECIFICITY: PRESENT IN ARTICULAR CHONDROCYTES, SYNOVIAL
CC       CELLS AS WELL AS IN LIVER. UNDETECTABLE IN MUSCLE TISSUES, LUNG,
CC       PANCREAS, MONONUCLEAR CELLS, OR FIBROBLASTS.
CC   -!- PTM: GLYCOSYLATED.
CC   -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
```

```
DR   EMBL; M80927; AAA16074.1; -.
DR   EMBL; Y08374; CAA69661.1; -.
DR   EMBL; Y08375; CAA69661.1; JOINED.
DR   EMBL; Y08376; CAA69661.1; JOINED.
DR   EMBL; Y08377; CAA69661.1; JOINED.
DR   EMBL; Y08378; CAA69661.1; JOINED.
DR   PIR; A49562; A49562.
DR   PDB; 1LA7; 10-APR-02.
DR   Genew; HGNC:1932; CHI3L1.
DR   MIM; 601525; -.
DR   GO; GO:0005578; C:extracellular matrix; TAS.
DR   GO; GO:0005615; C:extracellular space; TAS.
DR   GO; GO:0005207; F:extracellular matrix glycoprotein; TAS.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   InterPro; IPR001579; Glyco_hydro_18/2.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   ProDom; PD000471; Glyco_hydro_18; 1.
DR   SMART; SM00636; Glyco_18; 1.
DR   PROSITE; PS01095; CHITINASE_18; FALSE_NEG.
KW   Glycoprotein; Signal; 3D-structure.
FT   SIGNAL       1     21
FT   CHAIN       22    383       CHITINASE-3 LIKE PROTEIN 1.
FT   CARBOHYD    60     60       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   383 AA;  42513 MW;  76ADD8298EE8C2D1 CRC64;

  Query Match             100.0%;  Score 62;  DB 1;  Length 383;
  Best Local Similarity   100.0%;  Pred. No. 0.00023;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RSFTLASSETGVG 13
              |||||||||||||
Db        263 RSFTLASSETGVG 275


RESULT 2
C3L2_HUMAN
ID   C3L2_HUMAN     STANDARD;      PRT:   390 AA.
AC   Q15782; Q15749; Q15783;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Chitinase 3-like protein 2 precursor (YKL-39) (Chondrocyte protein
DE   39).
GN   CHI3L2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS LONG AND SHORT).
RA   Grossman A., Matsuyama T., Baker E., Waterhouse P., Sutherland G.R.,
RA   Mak T.W.;
RT   "Cloning of a novel lymphoid restricted human chitinase and
RT   localization to 1p13.3.";
RL   Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM SHORT).
RC   TISSUE=Articular cartilage;
RX   MEDLINE=96325055; PubMed=8702629;
RA   Hu B., Trinh K., Figueira W.F., Price P.A.;
RT   "Isolation and sequence of a novel human chondrocyte protein related
RT   to mammalian members of the chitinase protein family.";
RL   J. Biol. Chem. 271:19415-19420(1996).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=Short;
CC         IsoId=Q15782-1; Sequence=Displayed;
CC       Name=Long;
CC         IsoId=Q15782-2; Sequence=VSP_001066;
CC         Note=No experimental confirmation available;
CC   -!- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN CHONDROCYTES, FOLLOWED
CC       BY SYNOVIOCYTES, LUNG AND HEART. NOT DETECTED IN BRAIN, SPLEEN,
CC       PANCREAS, AND LIVER.
CC   -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U58515; AAB04534.1; -.
DR   EMBL; U58514; AAB04533.1; -.
DR   EMBL; U49835; AAC50597.1; ALT_INIT.
DR   Genew; HGNC:1933; CHI3L2.
DR   MIM; 601526; -.
DR   GO; GO:0005615; C:extracellular space; TAS.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   InterPro; IPR001579; Glyco_hydro_18/2.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   ProDom; PD000471; Glyco_hydro_18; 1.
DR   SMART; SM00636; Glyco_18; 1.
DR   PROSITE; PS01095; CHITINASE_18; FALSE_NEG.
KW   Hydrolase; Signal; Alternative splicing.
FT   SIGNAL       1     26       POTENTIAL.
FT   CHAIN       27    390       CHITINASE 3-LIKE PROTEIN 2.
FT   CARBOHYD    35     35       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC     1     12       MGATTMDQKSLW -> SQQLWLGSPDEVWLYLVCEHTHIFT
FT                                AIIWDSRTRFGSIDPSFHPSI (in isoform Long).
FT                                /FTId=VSP_001066.
SQ   SEQUENCE   390 AA;  43500 MW;  97B86A2F3AA35677 CRC64;

  Query Match             74.2%;  Score 46;  DB 1;  Length 390;
  Best Local Similarity   83.3%;  Pred. No. 0.29;
  Matches   10;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 SFTLASSETGVG 13
              |||||||:|| ||
Db        272 SFTLASAETTVG 283


RESULT 3
C3L1_MOUSE
ID   C3L1_MOUSE     STANDARD;      PRT:   381 AA.
AC   Q61362;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Chitinase-3 like protein 1 precursor (Cartilage glycoprotein-39)
DE   (GP-39) (BRP39 protein).
GN   CHI3L1 OR BRP39.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FVB/N; TISSUE=Breast;
RX   MEDLINE=95060797; PubMed=7970700;
RA   Morrison B.W., Leder P.;
RT   "neu and ras initiate murine mammary tumors that share genetic markers
RT   generally absent in c-myc and int-2-initiated tumors.";
RL   Oncogene 9:3417-3426(1994).
CC   -!- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN THE CAPACITY OF CELLS TO
CC       RESPOND TO AND COPE WITH CHANGES IN THEIR ENVIRONMENT (BY
CC       SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Extracellular (By similarity).
CC   -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
```

```
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X93035; CAA63603.1; -.
DR   PIR; S61551; S61551.
DR   MGD; MGI:1340899; Chi3l1.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   InterPro; IPR001579; Glyco_hydro_18/2.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   ProDom; PD000471; Glyco_hydro_18; 1.
DR   SMART; SM00636; Glyco_18; 1.
DR   PROSITE; PS01095; CHITINASE_18; FALSE_NEG.
KW   Glycoprotein; Signal.
FT   SIGNAL       1     21       POTENTIAL.
FT   CHAIN       22    381       CHITINASE-3 LIKE PROTEIN 1.
FT   CARBOHYD    60     60       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   381 AA;  43001 MW;  EF6581E8184F045C CRC64;

  Query Match           69.4%;  Score 43;  DB 1;  Length 381;
  Best Local Similarity   69.2%;  Pred. No. 1.1;
  Matches    9;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 RSFTLASSETGVG 13
              :|||||||| :|
Db        264 KSFTLASSENQLG 276


RESULT 4
RECN_MYCTU
ID   RECN_MYCTU     STANDARD;      PRT;   567 AA.
AC   O33197;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   DNA repair protein recN (Recombination protein N).
GN   RECN OR RV1696 OR MTI735 OR MTCI125.18.
OS   Mycobacterium tuberculosis.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX   NCBI_TaxID=1773;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=H37Rv;
RX   MEDLINE=98295987; PubMed=9634230;
RA   Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA   Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekaia F.,
RA   Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA   Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA   Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA   Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA   Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA   Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT   "Deciphering the biology of Mycobacterium tuberculosis from the
RT   complete genome sequence.";
RL   Nature 393:537-544(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CDC 1551 / Oshkosh;
RA   Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White C.,
RA   Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA   Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA   Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA   Bishai W.;
RT   "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT   laboratory strains.";
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: MAY BE INVOLVED IN RECOMBINATIONAL REPAIR OF DAMAGED
CC       DNA (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE RECN FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; Z98268; CAB10953.1; -.
DR   EMBL; AE007035; AAK46004.1; -.
DR   PIR; G70502; G70502.
DR   TIGR; MT1735; -.
DR   TubercuList; Rv1696; -.
DR   InterPro; IPR003439; ABC_transporter.
DR   InterPro; IPR004604; RecN.
DR   TIGRFAMs; TIGR00634; recN; 1.
KW   DNA repair; ATP-binding; Complete proteome.
FT   NP_BIND     29     36       ATP (POTENTIAL).
FT   CONFLICT   410    410       E -> G (IN REF. 2).
SQ   SEQUENCE   587 AA;  62228 MW;  4B7051EFB6C102E7 CRC64;

  Query Match           62.9%;  Score 39;  DB 1;  Length 587;
  Best Local Similarity   53.8%;  Pred. No. 11;
  Matches    7;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 RSFTLASSETGVG 13
              | ||: : ||| |
Db         22 RGFTVLTGETGTG 34


RESULT 5
K20A_MOUSE
ID   K20A_MOUSE     STANDARD;      PRT;   887 AA.
AC   P97329;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Kinesin family member 20A (Rabkinesin-6) (Rab6-interacting kinesin-
DE   like protein) (Kinesin-like protein 174).
GN   KIF20A OR RAB6KIFL.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/c; TISSUE=Testis;
RX   MEDLINE=98101856; PubMed=9438855;
RA   Echard A., Jollivet F., Martinez O., Lacapere J.-J., Rousselet A.,
RA   Janoueix-Lerosey I., Goud B.;
RT   "Interaction of a Golgi-associated kinesin-like protein with Rab6.";
RL   Science 279:580-585(1998).
CC   -!- FUNCTION: INTERACTS WITH GUANOSINE TRIPHOSPHATE (GTP)-BOUND FORMS
CC       OF RAB6A AND RAB6B. MAY ACT AS A MOTOR REQUIRED FOR THE RETROGRADE
CC       RAB6 REGULATED TRANSPORT OF GOLGI MEMBRANES AND ASSOCIATED
CC       VESICLES ALONG MICROTUBLES. HAS A MICROTUBULE PLUS END-DIRECTED
CC       MOTILITY.
CC   -!- SUBCELLULAR LOCATION: GOLGI.
CC   -!- TISSUE SPECIFICITY: UBIQUITOUSLY EXPRESSED, WITH HIGHEST LEVELS IN
CC       SPLEEN AND TESTIS.
CC   -!- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; Y09632; CAA70845.1; -.
DR   HSSP; P17119; 3KAR.
DR   MGD; MGI:1201682; Kif20a.
DR   InterPro; IPR001752; kinesin_motor.
DR   Pfam; PF00225; kinesin; 1.
DR   PRINTS; PR00380; KINESINHEAVY.
```

```
DR   SMART; SM00129; KISc; 1.
DR   PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; FALSE_NEG.
DR   PROSITE; PS50067; KINESIN_MOTOR_DOMAIN2; 1.
KW   Motor protein; Microtubules; ATP-binding; Coiled coil; Golgi stack;
KW   Protein transport; Transport.
FT   DOMAIN       69    494       KINESIN-MOTOR.
FT   NP_BIND     159    166       ATP (POTENTIAL).
FT   DOMAIN      559    804       COILED COIL (POTENTIAL).
FT   DOMAIN      805    887       GLOBULAR (POTENTIAL).
SQ   SEQUENCE   887 AA;  99876 MW;  7143CCB261A4EA15 CRC64;

  Query Match             61.3%;  Score 38;  DB 1;  Length 887;
  Best Local Similarity   66.7%;  Pred. No. 27;
  Matches    8;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 SFTLASSETGVG 13
              || | |:| |:|
Db        103 SFALKSNERGVG 114


RESULT 6
K20A_HUMAN
ID   K20A_HUMAN     STANDARD;      PRT:   890 AA.
AC   O95235;
DT   30-MAY-2000 (Rel. 38, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Kinesin family member 20A (Rabkinesin-6) (Rab6-interacting kinesin-
DE   like protein) (GG1C_2).
GN   KIF20A OR RAB6KIFL.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Endothelial cells;
RX   MEDLINE=99252096; PubMed=10233894;
RA   Horrevoets A.J.G., Fontijn R.D., van Zonneveld A.J., de Vries C.J.M.,
RA   ten Cate J.W., Pannekoek H.;
RT   "Vascular endothelial genes that are responsive to tumor necrosis
RT   factor-alpha in vitro are expressed in atherosclerotic lesions,
RT   including inhibitor of apoptosis protein-1, stannin, and two novel
RT   genes.";
RL   Blood 93:3418-3431(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=20267854; PubMed=10806357;
RA   Lai F., Fernald A.A., Zhao N., Le Beau M.M.;
RT   "cDNA cloning, expression pattern, genomic structure and chromosomal
RT   location of RAB6KIFL, a human kinesin-like gene.";
RL   Gene 248:117-125(2000).
CC   -!- FUNCTION: INTERACTS WITH GUANOSINE TRIPHOSPHATE (GTP)-BOUND FORMS
CC       OF RAB6A AND RAB6B. MAY ACT AS A MOTOR REQUIRED FOR THE RETROGRADE
CC       RAB6 REGULATED TRANSPORT OF GOLGI MEMBRANES AND ASSOCIATED
CC       VESICLES ALONG MICROTUBULES. HAS A MICROTUBULE PLUS END-DIRECTED
CC       MOTILITY (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Golgi (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF070672; AAC83230.1; -.
DR   EMBL; AF153329; AAD37806.1; -.
DR   HSSP; P17119; 3KAR.
DR   Genew; HGNC:9787; KIF20A.
DR   MIM; 605664; -.
DR   GO; GO:0005794; C:Golgi apparatus; TAS.
DR   GO; GO:0005488; F:vesicle transport; TAS.
DR   GO; GO:0016192; P:vesicle-mediated transport; TAS.
DR   InterPro; IPR001752; kinesin_motor.
DR   Pfam; PF00225; kinesin; 1.
DR   PRINTS; PR00380; KINESINHEAVY.
DR   SMART; SM00129; KISc; 1.
DR   PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; FALSE_NEG.
DR   PROSITE; PS50067; KINESIN_MOTOR_DOMAIN2; 1.
KW   Motor protein; Microtubules; ATP-binding; Coiled coil; Golgi stack;
KW   Protein transport; Transport.
FT   DOMAIN       70    495       KINESIN-MOTOR.
FT   NP_BIND     160    167       ATP (POTENTIAL).
FT   DOMAIN      611    762       COILED COIL (POTENTIAL).
FT   DOMAIN      763    890       GLOBULAR (POTENTIAL).
SQ   SEQUENCE   890 AA;  100277 MW;  6620264615496051 CRC64;

  Query Match             59.7%;  Score 37;  DB 1;  Length 890;
  Best Local Similarity   58.3%;  Pred. No. 42;
  Matches    7;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 SFTLASSETGVG 13
              || | |:| |:|
Db        104 SFALKSNERGIG 115


RESULT 7
GCP4_DROME
ID   GCP4_DROME     STANDARD;      PRT:   650 AA.
AC   Q9VKU7; Q9U1J7;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Gamma-tubulin complex component 4 homolog (Gamma ring complex protein
DE   75) (dGrip75).
GN   GRIP75 OR 75P OR CG6176.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20029771; PubMed=10562286;
RA   Fava F., Raynaud-Messina B., Leung-Tack J., Mazzolini L., Li M.,
RA   Guillemot J.C., Cachot D., Tollon Y., Ferrara P., Wright M.;
RT   "Human 76p: a new protein member of the gamma-tubulin associated
RT   protein family.";
RL   J. Cell Biol. 147:857-868(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
```

```
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [3]
RP   REVISIONS.
RC   STRAIN=Berkeley;
RX   MEDLINE=22425069; PubMed=12537572;
RA   Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA   Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA   Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA   Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA   Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA   Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA   Lewis S.E.;
RT   "Annotation of the Drosophila melanogaster euchromatic genome: a
RT   systematic review.";
RL   Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
CC   -!- FUNCTION: Gamma-tubulin complex is necessary for microtubule
CC       nucleation at the centrosome (By similarity).
CC   -!- SUBCELLULAR LOCATION: Centrosome (Potential).
CC   -!- SIMILARITY: BELONGS TO THE GCP FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE003628; AAF52962.2; -.
DR   EMBL; AJ249678; CAB62508.1; -.
DR   FlyBase; FBgn0026431; Grip75.
DR   GO; GO:0000274; C:gamma-tubulin ring complex; ISS.
DR   GO; GO:0005200; F:structural constituent of cytoskeleton; ISS.
DR   GO; GO:0007020; P:microtubule nucleation; ISS.
DR   Pfam; PF04130; Spc97_Spc98; 2.
KW   Microtubules.
SQ   SEQUENCE   650 AA;  74979 MW;  515D8B36A80D085E CRC64;

  Query Match           58.1%;  Score 36;  DB 1;  Length 650;
  Best Local Similarity  58.3%;  Pred. No. 46;
  Matches    7;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 RSFILASSETGV 12
              |:| ||:: |:|:
Db        386 RSFELAAFVTGI 397


RESULT 8
YO34_BPHP1
ID   YO34_BPHP1     STANDARD;      PRT;   187 AA.
AC   P51738;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical 20.8 kDa protein in LYS 3'region (ORF34).
OS   Bacteriophage HP1.
OC   Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae.
OX   NCBI_TaxID=10690;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HP1C1;
RX   MEDLINE=96279738; PubMed=8710508;
RA   Esposito D., Fitzmaurice W.P., Benjamin R.C., Goodman S.D.,
RA   Waldman A.S., Scocca J.J.;
RT   "The complete nucleotide sequence of bacteriophage HP1 DNA.";
RL   Nucleic Acids Res. 24:2360-2368(1996).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U24159; AAB09221.1; -.
DR   PIR; S69542; S69542.
KW   Hypothetical protein.
SQ   SEQUENCE   187 AA;  20761 MW;  A9A073D71DDEC02E CRC64;

  Query Match           56.5%;  Score 35;  DB 1;  Length 187;
  Best Local Similarity  63.6%;  Pred. No. 17;
  Matches    7;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          3 PTLASSETGVG 13
              | || :||| !
Db         72 PNLAEAETGKG 82


RESULT 9
TRPD_RHILO
ID   TRPD_RHILO     STANDARD;      PRT;   336 AA.
AC   Q98KE4;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Anthranilate phosphoribosyltransferase (EC 2.4.2.18).
GN   TRPD OR MLR0614.
OS   Rhizobium loti (Mesorhizobium loti).
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC   Phyllobacteriaceae; Mesorhizobium.
OX   NCBI_TaxID=381;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MAFF303099;
RX   MEDLINE=21082930; PubMed=11214968;
RA   Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA   Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA   Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA   Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA   Takeuchi C., Yamada M., Tabata S.;
RT   "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT   Mesorhizobium loti.";
RL   DNA Res. 7:331-338(2000).
CC   -!- CATALYTIC ACTIVITY: Anthranilate + phosphoribosylpyrophosphate =
CC       N-5'-phosphoribosyl-anthranilate + diphosphate.
CC   -!- PATHWAY: Tryptophan biosynthesis; second step.
CC   -!- SIMILARITY: BELONGS TO THE ANTHRANILATE PHOSPHORIBOSYLTRANSFERASE
CC       FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
```

```
DR   EMBL; AP002995; BAB48169.1; -.
DR   HAMAP; MF_00211; -; 1.
DR   InterPro; IPR005940; Ant_phspho_trans.
DR   InterPro; IPR000312; Glyco_trans_3.
DR   Pfam; PF02885; Glycos_trans_3N; 1.
DR   Pfam; PF00591; Glycos_transf_3; 1.
DR   ProDom; PD001864; Glyco_trans_3; 1.
DR   TIGRFAMs; TIGR01245; trpD; 1.
KW   Tryptophan biosynthesis; Transferase; Glycosyltransferase;
KW   Complete proteome.
SQ   SEQUENCE   336 AA;  34608 MW;  B7ABE45FC971D562 CRC64;

  Query Match             56.5%;  Score 35;  DB 1;  Length 336;
  Best Local Similarity   58.3%;  Pred. No. 34;
  Matches    7;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 RSFTLASSETGV 12
              |||:.   | :
Db        244 RSFTLTPEFVGL 255


RESULT 10
ALR_RICCN
ID   ALR_RICCN      STANDARD;      PRT;   356 AA.
AC   Q92JD9;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Alanine racemase (EC 5.1.1.1).
GN   ALR OR RC0128.
OS   Rickettsia conorii.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC   Rickettsiaceae; Rickettsieae; Rickettsia.
OX   NCBI_TaxID=781;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Malish 7;
RX   MEDLINE=21442074; PubMed=11557893;
RA   Ogata H., Audic S., Renesto-Audiffren P., Fournier P.-E., Barbe V.,
RA   Samson D., Roux V., Cossart P., Weissenbach J., Claverie J.-M.,
RA   Raoult D.;
RT   "Mechanisms of evolution in Rickettsia conorii and R. prowazekii.";
RL   Science 293:2093-2098(2001).
CC   -!- FUNCTION: Provides the D-alanine required for cell wall
CC       biosynthesis (By similarity).
CC   -!- CATALYTIC ACTIVITY: L-alanine = D-alanine.
CC   -!- COFACTOR: Pyridoxal phosphate (By similarity).
CC   -!- PATHWAY: D-alanine branch of peptidoglycan biosynthesis; first
CC       step.
CC   -!- SIMILARITY: Belongs to the alanine racemase family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AE008578; AAL02666.1; -.
DR   PIR; H97715; H97715.
DR   HAMAP; MF_01201; -; 1.
DR   InterPro; IPR000821; Ala_racemase.
DR   Pfam; PF00842; Ala_racemase_C; 1.
DR   Pfam; PF01168; Ala_racemase_N; 1.
DR   PRINTS; PR00992; ALARACEMASE.
DR   TIGRFAMs; TIGR00492; alr; 1.
DR   PROSITE; PS00395; ALANINE_RACEMASE; 1.
KW   Isomerase; Pyridoxal phosphate; Cell wall; Peptidoglycan synthesis;
KW   Complete proteome.
FT   ACT_SITE     34     34       CATALYTIC BASE SPECIFIC TO D-ALANINE
FT                                (BY SIMILARITY).
FT   ACT_SITE    249    249       CATALYTIC BASE SPECIFIC TO L-ALANINE
FT                                (BY SIMILARITY).
FT   BINDING      34     34       PYRIDOXAL PHOSPHATE (BY SIMILARITY).
SQ   SEQUENCE   355 AA;  39488 MW;  63A489ED384081EB CRC64;

  Query Match             56.5%;  Score 35;  DB 1;  Length 355;
  Best Local Similarity   66.7%;  Pred. No. 36;
  Matches    8;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 RSFTLASSETGV 12
              | | :|||| ||
Db         55 RHFFVASSEEGV 66


RESULT 11
MRP_MYCLE
ID   MRP_MYCLE      STANDARD;      PRT;   383 AA.
AC   P53382;
DT   01-OCT-1996 (Rel. 34, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Mrp protein homolog.
GN   MRP OR ML1080.
OS   Mycobacterium leprae.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX   NCBI_TaxID=1769;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Smith D.R., Robison K.;
RL   Submitted (MAY-1995) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TN;
RX   MEDLINE=21128732; PubMed=11234002;
RA   Cole S.T., Eiglmeier K., Parkhill J., James K.D., Thomson N.R.,
RA   Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
RA   Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
RA   Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,
RA   Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
RA   Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA   Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA   Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA   Barrell B.G.;
RT   "Massive gene decay in the leprosy bacillus.";
RL   Nature 409:1007-1011(2001).
CC   -!- FUNCTION: NOT KNOWN.
CC   -!- SIMILARITY: BELONGS TO THE MRP/NBP35 FAMILY OF ATP-BINDING
CC       PROTEINS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; U15180; AAA62909.1; ALT_INIT.
DR   EMBL; AL583920; CAC31461.1; -.
DR   PIR; B87044; B87044.
DR   Leproma; ML1080; -.
DR   InterPro; IPR002744; DUF59.
DR   InterPro; IPR000808; Mrp_family.
DR   Pfam; PF01883; DUF59; 1.
DR   ProDom; PD005595; DUF59; 1.
DR   PROSITE; PS01215; MRP; 1.
KW   ATP-binding; Complete proteome.
FT   NP_BIND     127    134       ATP (POTENTIAL).
SQ   SEQUENCE   383 AA;  40544 MW;  4D5B8F23A0C8279E CRC64;

  Query Match             56.5%;  Score 35;  DB 1;  Length 383;
  Best Local Similarity   53.8%;  Pred. No. 39;
  Matches    7;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;
```

```
Qy          1 RSFTLASSETGVG 13
              | :|:|:' : |||
Db        120 RVYTVASGKGGVG 132


RESULT 12
CN7B_SCHPO
ID   CN7B_SCHPO     STANDARD;      PRT;   402 AA.
AC   Q09722; Q9US29;
DT   01-NOV-1995 (Rel. 32, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   COP9/signalosome complex subunit 7B.
GN   CSN7B OR SPAC1751.03 OR SPAC31A2.01.
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
OC   Schizosaccharomyces.
OX   NCBI_TaxID=4896;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972;
RX   MEDLINE=21848401; PubMed=11859360;
RA   Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA   Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA   Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA   Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA   Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA   Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA   James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA   Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA   Oliver K., O'Neil S., Pearson D., Quail M.A., Rabbinowitsch E.,
RA   Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA   Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA   Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA   Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA   Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA   Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,
RA   Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA   Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA   Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA   Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA   Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA   Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA   Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA   Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA   Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT   "The genome sequence of Schizosaccharomyces pombe.";
RL   Nature 415:871-880(2002).
CC   -!- SIMILARITY: Contains 1 PCI domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AL133154; CAB61449.1; -.
DR   EMBL; Z50113; CAA90459.1; -.
DR   GeneDB_SPombe; SPAC1751.03; -.
DR   InterPro; IPR000717; PCI.
DR   Pfam; PF01399; PCI; 1.
DR   SMART; SM00088; PINT; 1.
KW   Signalosome.
FT   DOMAIN      274    365       PCI.
SQ   SEQUENCE   402 AA;  45085 MW;  CD972F781BFC0A3C CRC64;

  Query Match             55.5%;  Score 35;  DB 1;  Length 402;
  Best Local Similarity   70.0%;  Pred. No. 42;
  Matches    7;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 SFTLASSETG 11
              |||:|||: |
Db        382 SFTIASSKKG 391


RESULT 13
FLGL_VIBPA
ID   FLGL_VIBPA     STANDARD;      PRT;   445 AA.
AC   Q03475;
DT   01-OCT-1993 (Rel. 27, Created)
DT   15-SEP-2003 (Rel. 42, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Lateral flagellar hook-associated protein 2 (HAP2) (Filament CAP
DE   protein) (Flagellar cap protein).
GN   FLIDL OR LAFB OR VPA1550.
OS   Vibrio parahaemolyticus.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC   Vibrionaceae; Vibrio.
OX   NCBI_TaxID=670;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BB22;
RX   MEDLINE=93273702; PubMed=8501040;
RA   McCarter L.L., Wright M.E.;
RT   "Identification of genes encoding components of the swarmer cell
RT   flagellar motor and propeller and a sigma factor controlling
RT   differentiation of Vibrio parahaemolyticus.";
RL   J. Bacteriol. 175:3361-3371(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=RIMD 2210633 / Serotype O3:K6;
RX   MEDLINE=22508454; PubMed=12620739;
RA   Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
RA   Iijima Y., Najima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,
RA   Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;
RT   "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
RT   distinct from that of V. cholerae.";
RL   Lancet 361:743-749(2003).
CC   -!- FUNCTION: REQUIRED FOR THE MORPHOGENESIS AND FOR THE ELONGATION OF
CC       THE FLAGELLAR FILAMENT BY FACILITATING POLYMERIZATION OF THE
CC       FLAGELLIN MONOMERS AT THE TIP OF GROWING FILAMENT. FORMS A CAPPING
CC       STRUCTURE, WHICH PREVENTS FLAGELLIN SUBUNITS (TRANSPORTED THROUGH
CC       THE CENTRAL CHANNEL OF THE FLAGELLUM) FROM LEAKING OUT WITHOUT
CC       POLYMERIZATION AT THE DISTAL END. ESSENTIAL FOR SWARMING MOTILITY.
CC   -!- SUBUNIT: Homopentamer (By similarity).
CC   -!- SUBCELLULAR LOCATION: Flagellar.
CC   -!- MISCELLANEOUS: V.PARAHAEMOLYTICUS POSSESSES TWO FLAGELLAR SYSTEMS:
CC       A SINGLE POLAR FLAGELLUM PROPELS THE BACTERIUM IN LIQUID
CC       (SWIMMING), WHILE MULTIPLE LATERAL (PERITRICHOUS) FLAGELLA MOVE
CC       THE BACTERIUM OVER SURFACES (SWARMING). THE POLAR FLAGELLUM IS
CC       SYNTHESIZED CONSTITUTIVELY BUT LATERAL FLAGELLA ARE PRODUCED ONLY
CC       UNDER CONDITIONS IN WHICH THE POLAR FLAGELLUM IS NOT FUNCTIONAL.
CC   -!- SIMILARITY: BELONGS TO THE FLIC FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L06176; AAA27530.1; -.
DR   EMBL; U52957; AAB07351.1; -.
DR   EMBL; AP005089; BAC62893.1; -.
DR   PIR; B40590; B40590.
DR   InterPro; IPR003481; FliD.
DR   Pfam; PF02465; FliD; 1.
KW   Flagella; Coiled coil; Complete proteome.
FT   DOMAIN      388    423       COILED COIL (POTENTIAL).
FT   CONFLICT    226    226       K -> N (IN REF. 1).
SQ   SEQUENCE   445 AA;  48279 MW;  7B1836FC8722F00A CRC64;
```

```
  Query Match             56.5%;  Score 35;  DB 1;  Length 445;
  Best Local Similarity   70.0%;  Pred. No. 47;
  Matches    7;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 FTLASSETGV 12
              | :|:|:|||
Db        177 FMLSSIETGV 186


RESULT 14
RECN_TREPA
ID   RECN_TREPA     STANDARD;      PRT;   573 AA.
AC   O83456;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   DNA repair protein recN (Recombination protein N).
GN   RECN OR TP0442.
OS   Treponema pallidum.
OC   Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.
OX   NCBI_TaxID=160;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Nichols;
RX   MEDLINE=98332770; PubMed=9665876;
RA   Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
RA   Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,
RA   Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
RA   Khalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T.,
RA   McDonald L., Artiach P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA   Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA   Venter J.C.;
RT   "Complete genome sequence of Treponema pallidum, the syphilis
RT   spirochete.";
RL   Science 281:375-388(1998).
CC   -!- FUNCTION: MAY BE INVOLVED IN RECOMBINATIONAL REPAIR OF DAMAGED
CC       DNA (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE RECN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE001221; AAC65429.1; -.
DR   PIR; B71324; B71324.
DR   TIGR; TP0442; -.
DR   InterPro; IPR003439; ABC_transporter.
DR   InterPro; IPR004604; RecN.
DR   TIGRFAMs; TIGR00634; recN; 1.
KW   DNA repair; ATP-binding; Complete proteome.
FT   NP_BIND      29     36       ATP (POTENTIAL).
SQ   SEQUENCE   573 AA;  63571 MW;  9430825B90192683 CRC64;

  Query Match             56.5%;  Score 35;  DB 1;  Length 573;
  Best Local Similarity   63.6%;  Pred. No. 62;
  Matches    7;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          3 FTLASSETGVG 13
              || | ||| :
Db         24 FTALSGETGAG 34


RESULT 15
HMCT_HELPJ
ID   HMCT_HELPJ     STANDARD;      PRT;   686 AA.
AC   Q9ZL53;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Cadmium, zinc and cobalt transporting ATPase (EC 3.6.3.3)
DE   (EC 3.6.3.5).
GN   CADA OR JHP0727.
OS   Helicobacter pylori J99 (Campylobacter pylori J99).
OC   Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC   Helicobacteraceae; Helicobacter.
OX   NCBI_TaxID=85963;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99120557; PubMed=9923682;
RA   Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA   Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA   Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA   Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA   Trust T.J.;
RT   "Genomic sequence comparison of two unrelated isolates of the human
RT   gastric pathogen Helicobacter pylori.";
RL   Nature 397:176-180(1999).
CC   -!- FUNCTION: COUPLES THE HYDROLYSIS OF ATP WITH THE TRANSPORT OF
CC       CADMIUM, ZINC AND COBALT OUT OF THE CELL.
CC   -!- CATALYTIC ACTIVITY: ATP + H(2)O + Cd(2+)(In) = ADP + phosphate +
CC       Cd(2+)(Out).
CC   -!- CATALYTIC ACTIVITY: ATP + H(2)O + Zn(2+)(In) = ADP + phosphate +
CC       Zn(2+)(Out).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: Belongs to the cation transport ATPases family (P-type
CC       ATPases). Subfamily IB.
CC   -!- SIMILARITY: Contains 1 HMA domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE001504; AAD06308.1; -.
DR   PIR; E71895; E71895.
DR   InterPro; IPR006416; ATPase-1B_hvy.
DR   InterPro; IPR001757; ATPase_E1-E2.
DR   InterPro; IPR001366; Cad_ATPase.
DR   InterPro; IPR006404; Heavy_met_ATPase.
DR   InterPro; IPR006121; HeavyMe_transpt.
DR   InterPro; IPR005834; Hydrolase.
DR   InterPro; IPR006191; Metal_bind.
DR   Pfam; PF00122; E1-E2_ATPase; 1.
DR   Pfam; PF00403; HMA; 1.
DR   Pfam; PF00702; Hydrolase; 1.
DR   PRINTS; PR00119; CATATPASE.
DR   PRINTS; PR00941; CDATPASE.
DR   TIGRFAMs; TIGR01512; ATPase-IB2_Cd; 1.
DR   TIGRFAMs; TIGR01525; ATPase-IB_hvy; 1.
DR   TIGRFAMs; TIGR01494; ATPase_P-type; 3.
DR   PROSITE; PS00154; ATPASE_E1_E2; 1.
DR   PROSITE; PS01047; HMA_1; 1.
DR   PROSITE; PS50846; HMA_2; 1.
KW   Hydrolase; Transmembrane; Phosphorylation; Magnesium; ATP-binding;
KW   Metal-binding; Cadmium; Zinc; Cobalt; Complete proteome.
FT   DOMAIN       1     72       CYTOPLASMIC (BY SIMILARITY).
FT   TRANSMEM     73    92       1 (BY SIMILARITY).
FT   DOMAIN       93    102      EXTRACELLULAR (BY SIMILARITY).
FT   TRANSMEM     103   124      2 (BY SIMILARITY).
FT   DOMAIN       125   131      CYTOPLASMIC (BY SIMILARITY).
FT   TRANSMEM     132   151      3 (BY SIMILARITY).
FT   DOMAIN       152   154      EXTRACELLULAR (BY SIMILARITY).
FT   TRANSMEM     155   174      4 (BY SIMILARITY).
FT   DOMAIN       175   308      CYTOPLASMIC (BY SIMILARITY).
FT   TRANSMEM     309   327      5 (BY SIMILARITY).
FT   DOMAIN       328   332      EXTRACELLULAR (BY SIMILARITY).
FT   TRANSMEM     333   350      6 (BY SIMILARITY).
FT   DOMAIN       351   635      CYTOPLASMIC (BY SIMILARITY).
```

```
FT   TRANSMEM    636    657       7 (BY SIMILARITY).
FT   DOMAIN      658    665       EXTRACELLULAR (BY SIMILARITY).
FT   TRANSMEM    666    681       8 (BY SIMILARITY).
FT   DOMAIN      682    686       CYTOPLASMIC (BY SIMILARITY).
FT   DOMAIN        2     62       HMA.
FT   MOD_RES     368    368       PHOSPHORYLATION (BY SIMILARITY).
FT   METAL        11     11       POTENTIAL.
FT   METAL        14     14       POTENTIAL.
FT   METAL       583    583       MAGNESIUM (BY SIMILARITY).
FT   METAL       587    587       MAGNESIUM (BY SIMILARITY).
SQ   SEQUENCE   686 AA;  74837 MW;  C1E39A3B05747EE2 CRC64;

  Query Match             56.5%;  Score 35;  DB 1;  Length 686;
  Best Local Similarity   60.0%;  Pred. No. 76;
  Matches    6;  Conservative    3;  Mismatches    1;  Indels    3;  Gaps    2;

Qy          4 TLASETGVG 13
              |:|:|:: |:|
Db        590 TLASADVGIG 599


Search completed: September 24, 2003, 13:42:28
Job time : 9.17978 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 24, 2003, 13:33:45 ; Search time 42.2135 Seconds
(without alignments)
79.469 Million cell updates/sec

Title: US-09-744-282-6
Perfect score: 62
Sequence: 1 RSPTLASSETGVG 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_2virus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	62	100.0	383	4	Q96HI7	Q96hi7 homo sapien
2	62	100.0	383	4	Q8IVA4	Q8iva4 homo sapien
3	51	82.3	383	6	Q8SPQ0	Q8spq0 capra hircu
4	51	82.3	383	6	Q29411	Q29411 sus scrofa
5	48	77.4	332	6	O18949	O18949 bos taurus
6	46	74.2	352	11	Q9WTV1	Q9wtv1 rattus norv
7	46	74.2	390	4	Q96F97	Q96f97 homo sapien
8	43.5	70.2	396	11	Q9D7Q1	Q9d7q1 mus musculu
9	43	69.4	381	11	Q99J84	Q99j84 mus musculu
10	43	69.4	389	11	Q8BKL8	Q8bkl8 mus musculu
11	41	66.1	807	10	Q9XED2	Q9xed2 triticum ae
12	41	66.1	829	10	Q9XED3	Q9xed3 aegilops ta
13	41	66.1	830	10	O04074	O04074 triticum ae
14	40.5	65.3	387	4	Q9H3V8	Q9h3v8 homo sapien
15	40.5	65.3	466	4	Q13231	Q13231 homo sapien
16	40	64.5	185	9	Q38137	Q38137 bacteriopha
17	40	64.5	467	16	Q8XAI5	Q8xai5 escherichia
18	40	64.5	486	16	Q8FJ25	Q8fj25 escherichia
19	39	62.9	896	18	Q8DH74	Q8dh74 synechococc
20	38	61.3	261	17	Q9HPA0	Q9hpa0 halobacteri
21	38	61.3	286	16	Q8U941	Q8u941 agrobacteri
22	38	61.3	441	2	O52483	O52483 serratia ma
23	38	61.3	531	10	Q9SRK9	Q9srk9 arabidopsis
24	38	61.3	563	10	Q8W4N1	Q8w4n1 arabidopsis
25	38	61.3	565	16	Q8XJE4	Q8xje4 clostridium
26	38	61.3	1593	5	Q9V8J3	Q9v8j3 drosophila
27	37	59.7	132	16	Q8RG37	Q8rg37 fusobacteri
28	37	59.7	199	11	Q8C5D6	Q8c5d6 mus musculu
29	37	59.7	242	17	O27260	O27260 methanobact
30	37	59.7	298	16	Q92W69	Q92w69 rhizobium m
31	37	59.7	315	16	Q8FHF6	Q8fhf6 escherichia
32	37	59.7	459	16	Q9KBG2	Q9kbg2 bacillus ha
33	37	59.7	793	12	Q99CY1	Q99cy1 bovine herp
34	37	59.7	931	5	Q81S90	Q81s90 leishmania
35	36	58.1	203	13	Q8JG24	Q8jg24 kassina mac
36	36	58.1	251	17	O27403	O27403 methanobact
37	36	58.1	266	16	O69626	O69626 mycobacteri
38	36	58.1	278	2	Q50202	Q50202 mycobacteri
39	36	58.1	298	2	Q9Z406	Q9z406 pseudomonas
40	36	58.1	351	15	Q9CCX7	Q9ccx7 mycobacteri
41	36	58.1	368	16	Q8F072	Q8f072 leptospira
42	36	58.1	396	5	Q8MVU1	Q8mvu1 homarus gam
43	36	58.1	424	16	Q9PPC8	Q9ppc8 campylobact
44	36	58.1	446	10	Q6SXN2	Q6sxn2 pyrus pyrif
45	36	58.1	446	10	Q8S3N8	Q8s3n8 oryza sativ

ALIGNMENTS

RESULT 1
Q96HI7
ID Q96HI7 PRELIMINARY; PRT; 383 AA.
AC Q96HI7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Similar to chitinase 3-like 1 (Cartilage glycoprotein-39).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC008568; AAH08568.1; -.
DR InterPro; IPR001223; Glyco_hydro_18.
DR InterPro; IPR005829; Sug_transporter.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR SMART; SM00636; Glyco_18; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
SQ SEQUENCE 383 AA; 42625 MW; 6C9EF133BCC7C2D1 CRC64;

Query Match 100.0%; Score 62; DB 4; Length 383;
Best Local Similarity 100.0%; Pred. No. 0.0016;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RSPTLASSETGVG 13
|||||||||||||
Db 263 RSPTLASSETGVG 275

RESULT 2
Q8IVA4
ID Q8IVA4 PRELIMINARY; PRT; 383 AA.
AC Q8IVA4;

```
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Similar to chitinase 3-like 1 (Cartilage glycoprotein-39).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RA   Strausberg R.;
RL   Submitted (OCT-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC038354; AAH38354.1; -.
SQ   SEQUENCE   383 AA;  42526 MW;  76345433BDC7C6C0 CRC64;

  Query Match           100.0%;  Score 62;  DB 4;  Length 383;
  Best Local Similarity  100.0%;  Pred. No. 0.0016;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RSFTLASSETGVG 13
              |||||||||||||
Db        263 RSFTLASSETGVG 275


RESULT 3
Q8SPQ0
ID   Q8SPQ0      PRELIMINARY;      PRT;   383 AA.
AC   Q8SPQ0;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   BP40 precursor.
OS   Capra hircus (Goat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Caprinae; Capra.
OX   NCBI_TaxID=9925;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Mammary gland;
RA   Paramasivam M., Saravanan K., Sharma S., Mohanty A.K., Yadav S.,
RA   Srinivasan A., Singh T.P.;
RT   "Goat mammary gland mRNA for a novel protein, BP40, complete coding
RT   region.";
RL   Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY081150; AAL87007.1; -.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   ProDom; PD000471; Glyco_hydro_18; 1.
DR   SMART; SM00636; Glyco_18; 1.
KW   Signal.
FT   SIGNAL       1     16       POTENTIAL.
SQ   SEQUENCE   383 AA;  42893 MW;  17655ED4BB4E9F5F CRC64;

  Query Match           82.3%;  Score 51;  DB 6;  Length 383;
  Best Local Similarity  84.6%;  Pred. No. 0.21;
  Matches   11;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 RSFTLASSETGVG 13
              |||||||||:|.|
Db        263 RSFTLASSKTDVG 275


RESULT 4
Q29411
ID   Q29411      PRELIMINARY;      PRT;   383 AA.
AC   Q29411;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   38 kDa heparin-binding glycoprotein.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Smooth muscle;
RX   MEDLINE=95286589; PubMed=7768902;
RA   Shackelton L.M., Mann D.M., Millis A.J.;
RT   "Identification of a 38-kDa heparin-binding glycoprotein (gp38k) in
RT   differentiating vascular smooth muscle cells as a member of a group of
RT   proteins associated with tissue remodeling.";
RL   J. Biol. Chem. 270:13076-13083(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Smooth muscle;
RA   Shackelton L.M., Mann D.M., Millis A.J.F.;
RL   Submitted (JAN-1995) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U19900; AAA86482.1; -.
DR   EMBL; Z47803; CAA87764.1; -.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   ProDom; PD000471; Glyco_hydro_18; 1.
DR   SMART; SM00636; Glyco_18; 1.
KW   Heparin-binding.
SQ   SEQUENCE   383 AA;  42443 MW;  3D1039F49910BDC4 CRC64;

  Query Match           82.3%;  Score 51;  DB 6;  Length 383;
  Best Local Similarity  84.6%;  Pred. No. 0.21;
  Matches   11;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 RSFTLASSETGVG 13
              |||||||||:|.||
Db        263 RSFTLASSKTDVG 275


RESULT 5
O18949
ID   O18949      PRELIMINARY;      PRT;   332 AA.
AC   O18949;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Chitinase-like protein 1 (Fragment).
GN   CLP-1.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Recklies A.D., White C.;
RT   "Expression of chitinase-like protein 1 (CLP-1) in bovine
RT   chondrocytes.";
RL   Submitted (JUN-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF011373; AAB64304.1; -.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   ProDom; PD000471; Glyco_hydro_18; 1.
DR   SMART; SM00636; Glyco_18; 1.
FT   NON_TER       1       1
FT   NON_TER     332     332
SQ   SEQUENCE   332 AA;  37434 MW;  03F163323486408F CRC64;

  Query Match           77.4%;  Score 48;  DB 6;  Length 332;
  Best Local Similarity  76.9%;  Pred. No. 0.69;
  Matches   10;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 RSFTLASSETGVG 13
              ||:|||||:|.||
Db        239 RSYTLASSKTDVG 251
```

```
RESULT 6
Q9WTV1
ID   Q9WTV1      PRELIMINARY;      PRT:   352 AA.
AC   Q9WTV1;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Glycoprotein-39 (Fragment).
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Lewis;
RA   Wendling U., Boots A.M.H., van Eden W.;
RT   "Cloning of the rat homologue of Human Cartilage glycoprotein-39 a
RT   potential autoantigen in arthritis.";
RL   Submitted (APR-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF062038; AAD22610.1; -.
DR   HSSP; P07254; 1CTN.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   ProDom; PD000471; Glyco_hydro_18; 1.
DR   SMART; SM00636; Glyco_18; 1.
FT   NON_TER       1      1
FT   NON_TER     352    352
SQ   SEQUENCE   352 AA;  39391 MW;  CBDE991610AC936C CRC64;

  Query Match             74.2%;  Score 46;  DB 11;  Length 352;
  Best Local Similarity   76.9%;  Pred. No. 1.8;
  Matches   10;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 RSFTLASSETGVG 13
              :|||||:|:  |:
Db        242 KSFTLASSENQVG 254


RESULT 7
Q96F97
ID   Q96F97      PRELIMINARY;      PRT:   390 AA.
AC   Q96F97;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Chitinase 3-like 2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RA   Strausberg R.;
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC011460; AAH11460.1; -.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   ProDom; PD000471; Glyco_hydro_18; 1.
DR   SMART; SM00636; Glyco_18; 1.
SQ   SEQUENCE   390 AA;  43529 MW;  6801E64AF721AD83F CRC64;

  Query Match             74.2%;  Score 46;  DB 4;  Length 390;
  Best Local Similarity   83.3%;  Pred. No. 2;
  Matches   10;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 SFTLASSETGVG 13
              |||:|||:|:||
Db        272 SFTLASAETTVG 283


RESULT 8
Q9D7Q1
ID   Q9D7Q1      PRELIMINARY;      PRT:   396 AA.
AC   Q9D7Q1;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   2300002L19Rik protein.
GN   2300002L19RIK.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Tongue;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriquez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
DR   EMBL; AK009012; BAB26025.1; -.
DR   MGD; MGI:1919134; 2300002L19Rik.
DR   InterPro; IPR000677; 2S_Globulin.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   InterPro; IPR001579; Glyco_hydro_18/2.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   PRINTS; PR00551; 2SGLOBULIN.
DR   ProDom; PD000471; Glyco_hydro_18; 1.
DR   SMART; SM00636; Glyco_18; 1.
DR   PROSITE; PS01095; CHITINASE_18; 1.
SQ   SEQUENCE   396 AA;  43413 MW;  32BB674B81AF7123 CRC64;

  Query Match             70.2%;  Score 43.5;  DB 11;  Length 396;
  Best Local Similarity   78.6%;  Pred. No. 6.3;
  Matches   11;  Conservative    1;  Mismatches    1;  Indels    1;  Gaps    1;

Qy          1 RSFTLA-SSETGVG 13
              |:|||| ||: |||
Db        269 RSFTLASSSDNGVG 282


RESULT 9
Q99C84
ID   Q99C84      PRELIMINARY;      PRT:   381 AA.
AC   Q99C84;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Similar to chitinase 3-like 1 (Cartilage glycoprotein-39).
GN   CHI3L1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Strausberg R.;
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
```

```
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Strausberg R.;
RL   Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Strausberg R.;
RL   Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC005511; AAH05611.1; -.
DR   EMBL; BC003780; AAH03780.1; -.
DR   EMBL; BC004734; AAH04734.1; -.
DR   MGD; MGI:1340899; Chi3l1.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   ProDom; PD000471; Glyco_hydro_18; 1.
DR   SMART; SM00636; Glyco_18; 1.
SQ   SEQUENCE   381 AA;  42979 MW;  FF6588C5AE9D4450 CRC64;

  Query Match             69.4%;  Score 43;  DB 11;  Length 381;
  Best Local Similarity   69.2%;  Pred. No. 7.6;
  Matches    9;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 RSFTLASSETGVG 13
              :|||||||:  :|
Db        264 KSFTLASSENQLG 276


RESULT 10
Q8BKL8
ID   Q8BKL8      PRELIMINARY;      PRT;   389 AA.
AC   Q8BKL8;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Similar to chitinase 3-like 1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Dorsal root ganglion;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
DR   EMBL; AK051475; BAC34654.1; -.
SQ   SEQUENCE   389 AA;  43893 MW;  9E7D66069B233834 CRC64;

  Query Match             69.4%;  Score 43;  DB 11;  Length 389;
  Best Local Similarity   69.2%;  Pred. No. 7.8;
  Matches    9;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 RSFTLASSETGVG 13
              :|||||||:  :|
Db        272 KSFTLASSENQLG 284


RESULT 11
Q9XED2
ID   Q9XED2      PRELIMINARY;      PRT;   807 AA.
AC   Q9XED2;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Starch branching enzyme-I.
GN   SBE-I.
OS   Triticum aestivum (Wheat).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC   Triticeae; Triticum.
OX   NCBI_TaxID=4565;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Endosperm;
RA   Rahman S., Li Z., Abrahams S., Abbott D., Appels R., Morell M.K.;
RT   "Characterisation of a gene encoding wheat endosperm starch branching
RT   enzyme-I.";
RL   Theor. Appl. Genet. 98:156-163(1999).
DR   EMBL; AF076679; AAD30186.1; -.
DR   InterPro; IPR006047; Alpha_amyl_cat.
DR   InterPro; IPR004193; Glyco_hydro_13N.
DR   Pfam; PF00128; alpha-amylase; 1.
DR   Pfam; PF02922; isoamylase_N; 1.
SQ   SEQUENCE   807 AA;  91390 MW;  B9E27F0158371813 CRC64;

  Query Match             66.1%;  Score 41;  DB 10;  Length 807;
  Best Local Similarity   53.8%;  Pred. No. 44;
  Matches    7;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 RSFTLASSETGVG 13
              | :|:|::| |||
Db         43 RDYIMATAEDGVG 55


RESULT 12
Q9XED3
ID   Q9XED3      PRELIMINARY;      PRT;   829 AA.
AC   Q9XED3;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Starch branching enzyme-I.
GN   SBE-I.
OS   Aegilops tauschii (Aegilops squarrosa).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC   Triticeae; Aegilops.
OX   NCBI_TaxID=37682;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Rahman S., Li Z., Abrahams S., Abbott D., Appels R., Morell M.K.;
RT   "Characterisation of a gene encoding wheat endosperm starch branching
RT   enzyme-I.";
RL   Theor. Appl. Genet. 98:156-163(1999).
DR   EMBL; AF076680; AAD30187.1; -.
DR   InterPro; IPR006047; Alpha_amyl_cat.
DR   InterPro; IPR004193; Glyco_hydro_13N.
DR   Pfam; PF00128; alpha-amylase; 1.
DR   Pfam; PF02922; isoamylase_N; 1.
SQ   SEQUENCE   829 AA;  93667 MW;  0E5534988F4877C7 CRC64;

  Query Match             66.1%;  Score 41;  DB 10;  Length 829;
  Best Local Similarity   53.8%;  Pred. No. 45;
  Matches    7;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 RSFTLASSETGVG 13
              | :|:|::| |||
Db         66 RDYIMATAEDGVG 78


RESULT 13
O04074
ID   O04074      PRELIMINARY;      PRT;   830 AA.
AC   O04074;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Starch branching enzyme I precursor (Starch branching ENYZYME 1)
DE   (EC 2.4.1.81).
GN   SBE1 OR SBE1D.
OS   Triticum aestivum (Wheat).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
```

```
OC   Triticeae; Triticum.
OX   NCBI_TaxID=4565;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. FIELDER;
RA   Repellin A., Nair R., Baga M., Chibbar R.N.;
RL   Plant Physiol. 114:1145-1145(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Cheyenne; TISSUE=Endosperm;
RA   McCue K.F., Burkman W.J., Tanaka C.K., Anderson O.D.;
RT   "Starch Branching Enzymes SBE1 and SBE2 from Wheat (Triticum aestivum
RT   cv. Cheyenne): Molecular Characterization, Developmental Expression,
RT   and Homolog Assignment by Differential PCR.";
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; Y12320; CAA72987.1; -.
DR   EMBL; AF286317; AAG27621.1; -.
DR   InterPro; IPR006047; Alpha_amyl_cat.
DR   InterPro; IPR004193; Glyco_hydro_13N.
DR   Pfam; PF00128; alpha-amylase; 1.
DR   Pfam; PF02922; isoamylase_N; 1.
KW   Glycosyltransferase; Signal; Transferase.
FT   SIGNAL        1     61       POTENTIAL.
FT   CHAIN        62    830       POTENTIAL.
SQ   SEQUENCE   830 AA;  93577 MW;  3599218320FBA67F CRC64;

  Query Match             66.1%;  Score 41;  DB 10;  Length 830;
  Best Local Similarity   53.8%;  Pred. No. 45;
  Matches    7;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 RSFTLASSETGVG 13
              | :|:|:::| ||:
Db         66 RDYTMATAEDGVG 78


RESULT 14
Q9H3V8
ID   Q9H3V8      PRELIMINARY;      PRT;   387 AA.
AC   Q9H3V8;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Chitotriosidase precursor.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96064695; PubMed=7592832;
RA   Boot R.G., Renkema G.H., Strijland A., van Zonneveld A.J., Aerts J.M.;
RT   "Cloning of a cDNA encoding chitotriosidase, a human chitinase
RT   produced by macrophages.";
RL   J. Biol. Chem. 270:26252-26256(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98421482; PubMed=9748235;
RA   Boot R.G., Renkema G.H., Verhoek M., Strijland A., Bliek J.,
RA   de Meulemeester T.M., Mannens M.M., Aerts J.M.;
RT   "The human chitotriosidase gene. Nature of inherited enzyme
RT   deficiency.";
RL   J. Biol. Chem. 273:25680-25685(1998).
DR   EMBL; U62662; AAG10644.1; -.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   InterPro; IPR001579; Glyco_hydro_18/2.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   ProDom; PD000471; Glyco_hydro_18; 1.
DR   SMART; SM00636; Glyco_18; 1.
DR   PROSITE; PS01095; CHITINASE_18; 1.
KW   Signal.
FT   SIGNAL        1     21       POTENTIAL.
FT   CHAIN        22    387       CHITOTRIOSIDASE.
SQ   SEQUENCE   387 AA;  43133 MW;  03A272B8BC5E0D71 CRC64;

  Query Match             65.3%;  Score 40.5;  DB 4;  Length 387;
  Best Local Similarity   78.6%;  Pred. No. 24;
  Matches   11;  Conservative    1;  Mismatches    1;  Indels    1;  Gaps    1;

Qy          1 RSFTLA-SSETGVG 13
              |||||| ||:| ||
Db        269 RSFTLASSSDTRVG 282


RESULT 15
Q13231
ID   Q13231      PRELIMINARY;      PRT;   466 AA.
AC   Q13231;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Chitotriosidase precursor.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96064695; PubMed=7592832;
RA   Boot R.G., Renkema G.H., Strijland A., van Zonneveld A.J., Aerts J.M.;
RT   "Cloning of a cDNA encoding chitotriosidase, a human chitinase
RT   produced by macrophages.";
RL   J. Biol. Chem. 270:26252-26256(1995).
DR   EMBL; U29615; AAC50246.1; -.
DR   Genew; HGNC:1936; CHIT1.
DR   InterPro; IPR002557; Chitin_bind_PerA.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   InterPro; IPR001579; Glyco_hydro_18/2.
DR   Pfam; PF01607; CBM_14; 1.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   ProDom; PD000471; Glyco_hydro_18; 1.
DR   SMART; SM00494; ChtBD2; 1.
DR   SMART; SM00636; Glyco_18; 1.
DR   PROSITE; PS01095; CHITINASE_18; 1.
KW   Signal.
FT   SIGNAL        1     21       POTENTIAL.
FT   CHAIN        22    466       CHITOTRIOSIDASE.
SQ   SEQUENCE   466 AA;  51681 MW;  B4312D1E855E386D CRC64;

  Query Match             65.3%;  Score 40.5;  DB 4;  Length 466;
  Best Local Similarity   78.6%;  Pred. No. 29;
  Matches   11;  Conservative    1;  Mismatches    1;  Indels    1;  Gaps    1;

Qy          1 RSFTLA-SSETGVG 13
              |||||| ||:| ||
Db        269 RSFTLASSSDTRVG 282


Search completed: September 24, 2003, 13:47:26
Job time : 43.2135 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run On: September 24, 2003, 13:25:57 ; Search Time 52.1461 Seconds
(without alignments)
39.570 Million cell updates/sec

Title: US-09-744-282-7
Perfect score: 65
Sequence: 1 VGYDDQESVKSKV 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	65	100.0	13	17	AAR95077	Peptide contg. hC
2	65	100.0	13	21	AAY69232	Peptide fragment c
3	65	100.0	13	23	AAE20225	Human cartilage gl
4	65	100.0	105	22	AAU02074	Synthetic human ta
5	65	100.0	220	22	AAU02078	Synthetic human mu
6	65	100.0	281	22	AAU02079	Synthetic multi-ta
7	65	100.0	321	9	AAP81342	Polypeptide involv
8	65	100.0	350	22	AAU02076	Synthetic human mu
9	65	100.0	383	23	ABG96297	Human ovarian canc
10	65	100.0	383	24	ABG56651	Lung cancer-associ
11	65	100.0	459	22	AAU02077	Synthetic multi-ta
12	61	93.8	16	18	AAW45150	Human cartilage gl
13	58	89.2	383	21	AAB03442	Gp38k protein seqc
14	55	84.6	16	18	AAW45151	Human cartilage gl
15	52	80.0	14	22	ABB55896	Vascular dementia-
16	52	80.0	14	22	ABB55940	Vascular dementia-
17	52	80.0	14	22	ABB56115	Vascular dementia-
18	52	80.0	14	22	ABB52309	Human API-41 trypt
19	52	80.0	14	22	AAU28537	DP1 tryptic digest
20	52	80.0	14	22	AAU28750	DP1 tryptic digest
21	52	80.0	14	22	AAU26184	Depression-Associa
22	52	80.0	14	22	AAU26398	Depression-Associa
23	52	80.0	383	18	AAW26751	Bovine whey protei
24	47	72.3	739	22	ABB81049	N. tabacum chitin-
25	46	70.8	9	17	AAR95073	Human cartilage gl
26	46	70.8	9	21	AAY69228	Peptide fragment c
27	46	70.8	372	23	AAE25906	Human chitinase pr
28	46	70.8	373	19	AAW40262	Human chitinase pr
29	46	70.8	373	19	AAW40261	Human chitinase pr
30	46	70.8	373	20	AAY42427	Clone of the C-ter
31	46	70.8	373	20	AAY42426	Chitinase amino ac
32	46	70.8	373	23	AAE25905	Human chitinase-re
33	46	70.8	373	23	ABB76293	Human chitinase tr
34	46	70.8	373	23	ABB76294	Human chitinase an
35	46	70.8	387	18	AAW08585	Human 39 kDa chiti
36	46	70.8	421	23	ABB92991	Herbicidally activ
37	46	70.8	464	18	AAW31497	Human chitotriosid
38	46	70.8	466	18	AAW31498	Human chitotriosid
39	46	70.8	466	18	AAW08584	Human 50 kDa chiti
40	46	70.8	466	19	AAW40259	Human chitinase pr
41	46	70.8	466	19	AAW40260	Human chitinase pr
42	46	70.8	466	20	AAY42425	MO-218 clone of hu
43	46	70.8	466	20	AAY42426	MO-15B clone of hu
44	46	70.8	466	22	AAB00432	Human chitinase pr
45	46	70.8	466	22	AAB00433	Human chitinase pr

ALIGNMENTS

RESULT 1
AAR95077
ID AAR95077 standard; peptide; 13 AA.
XX
AC AAR95077;
XX
DT 17-DEC-1996 (first entry)
XX
DE Peptide contg. hC gp-39 autoantigen derived peptide (III).
XX
KW Human; cartilage; glycoprotein 39; HC gp-39; autoantigen;
KW Induction; T-cell; tolerance; mediated; destruction;
KW rheumatoid arthritis; disease model; diagnosis.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 3..11
FT /note= "HC gp-39 peptide (III)"
XX
PN WO9613517-A1.
XX
PD 09-MAY-1996.
XX
PF 25-OCT-1995; 95WO-EP04203.
XX
PR 07-APR-1995; 95EP-0200886.
PR 27-OCT-1994; 94EP-0203128.
XX
PA (ALKO) AKZO NOBEL NV.
XX
PI Boots AMH, Verheijden GFM;

XX
DR WPI; 1996-239447/24.
XX
PT Peptide(S) derived from human cartilage gp-39 autoantigen - used
PT to induce specific T cell tolerance, esp. for treatment of
PT rheumatoid arthritis, also useful in producing arthritic animal
PT models
XX
PS Claim 2; Page 28; 37pp; English.
XX
CC The present peptide, contg. a peptide derived from the human
CC cartilage glycoprotein 39 (hC gp-39) autoantigen, can be used to
CC induce specific T-cell tolerance to hC gp-39 in a patient with
CC T-cell mediated cartilage destruction, esp. rheumatoid arthritis
CC (RA). It may also be used to induce RA in mice, etc. (human disease
CC models), and to diagnose the presence of activated T-cells
CC autoreactive with hC gp-39, i.e. to identify patients suitable for
CC tolerising treatment.
CC Therapeutically the peptide is given in a dose of 0.1-100 microg/kg
CC by injection, or 10-50 microg/kg to produce a disease model. The
CC peptide has a specific tolerising effect on autoreactive T-cells,
CC but leaves other components of the immune system intact, and
CC therefore should have few side effects.
XX
SQ Sequence 13 AA;

Query Match 100.0%; Score 65; DB 17; Length 13;
Best Local Similarity 100.0%; Pred. No. 2.6e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VGYDDQESVKSKV 13
||||||||||||| |
Db 1 VGYDDQESVKSKV 13

RESULT 2
AAY69232
ID AAY69232 standard; peptide; 13 AA.
XX
AC AAY69232;
XX
DT 30-MAY-2000 (first entry)
XX
DE Peptide fragment of human cartilage gp-39 (HC gp-39) protein.
XX
KW Human cartilage gp-39 protein; HC gp-39 protein; T-cell activator;
KW inflammatory disease; lymphocyte; antigen; bystander suppression;
KW antigen-specific T-cell tolerance; autoimmune disorder;
KW inflammatory disease; rheumatoid arthritis;
KW autoimmune inflammatory disease; autoantigen.
XX
OS Homo sapiens.
XX
PN WO200004917-A2.
XX
PD 03-FEB-2000.
XX
PF 19-JUL-1999; 99WO-EP05331.
XX
PR 23-JUL-1998; 98EP-0202471.
XX
PA (ALKU) AKZO NOBEL NV.
XX
PI Miltenburg AMM, Boots AMH;
XX
DR WPI; 2000-182537/16.
XX
PT Use of cartilage protein or peptides for treating inflammatory diseases
PT especially rheumatoid arthritis -
XX
PS Claim 2; Page 17; 29pp; English.
XX
CC AAY69226-33 represent fragments of the human cartilage gp-39 (HC gp-39)
CC protein. The peptides are T-cell activators. The peptides are used for
CC preventing inflammatory diseases, by modulating lymphocytes reactive
CC to antigens other than HC gp-39, present in the same tissue of
CC HC gp-39. By the induction of antigen-specific T-cell tolerance,
CC autoimmune disorders can be treated by bystander suppression. The
CC HC gp-39 peptides are useful for preventing inflammatory diseases
CC especially rheumatoid arthritis. The fragments may also be used for
CC treating autoimmune inflammatory diseases. The HC gp-39 peptides are
CC especially beneficial in autoimmune conditions in which the
CC autoantigens associated with the condition are unknown.
XX
SQ Sequence 13 AA;

Query Match 100.0%; Score 65; DB 21; Length 13;
Best Local Similarity 100.0%; Pred. No. 2.6e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VGYDDQESVKSKV 13
||||||||||||| |
Db 1 VGYDDQESVKSKV 13

RESULT 3
AAE20225
ID AAE20225 standard; peptide; 13 AA.
XX
AC AAE20225;
XX
DT 18-JUN-2002 (first entry)
XX
DE Human cartilage glycoprotein (HC gp)-39 peptide (residues 326-338).
XX
KW Human cartilage glycoprotein-39; HC gp-39; autoimmune disorder; MHC;
KW major histocompatibility complex; insulin-dependent diabetes mellitus;
KW multiple sclerosis; psoriasis; myasthenia gravis; rheumatoid arthritis;
KW immunosuppressive; neuroprotective; therapy.
XX
OS Homo sapiens.
XX
PN WO200214870-A2.
XX
PD 21-FEB-2002.
XX
PF 08-AUG-2001; 2001WO-EP09136.
XX
PR 14-AUG-2000; 2000EP-0202844.
XX
PA (ALKU) AKZO NOBEL NV.
XX
PI Steenbakkers PGA;
XX
DR WPI; 2002-269214/31.
XX
PT Method for diagnosing an autoimmune disease activity by detecting the
PT presence of an autoimmune specific major histocompatibility complex
PT (MHC)-peptide complex in a patient using antibodies specific against
PT MHC-peptide complexes -
XX
PS Disclosure; Page 43; 43pp; English.
XX
CC The patent discloses a method for diagnosing an autoimmune disease
CC activity. The method involves detecting the presence of an autoimmune
CC specific major histocompatibility complex (MHC)-peptide complex in a
CC patient suffering from an autoimmune disease with antibodies or its
CC antigen binding domains that specifically bind to a MHC-gp39-derived
CC peptide. The method and antibodies are useful for diagnosing auto-
CC immune diseases or purifying autoimmune specific MHC-peptide complexes.
CC They are also useful in therapy or for the manufacture of pharmaceutical
CC preparation for the treatment of autoimmune disorders such as insulin-
CC dependent diabetes mellitus, multiple sclerosis, myasthenia gravis,
CC psoriasis or rheumatoid arthritis. The present sequence is human
CC cartilage glycoprotein (HC gp)-39 peptide (residues 326-338). This
CC peptide is used in the exemplification of the invention.

```
XX
SQ   Sequence   13 AA;

  Query Match             100.0%;  Score 65;  DB 23;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 2.6e-05;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VGYDDQESVKSKV 13
              |||||||||||||
Db          1 VGYDDQESVKSKV 13


RESULT 4
AAU02074
ID   AAU02074 standard; Protein; 105 AA.
XX
AC   AAU02074;
XX
DT   07-SEP-2001  (first entry)
XX
DE   Synthetic human target autoantigen shGLP/RA.
XX
KW   shGLP/RA; immunogenic epitope cluster; IEC;
KW   synthetic human chondrocyte glycoprotein 69/cartilage link protein;
KW   autoantigen; autoimmune disease; multiple sclerosis;
KW   rheumatoid arthritis; insulin-dependent diabetes mellitus;
KW   myasthenia gravis; uveitis; autoimmune hepatitis; thyroiditis; insulitis;
KW   orchitis; idiopathic thrombocytopenic purpura; inflammatory disease;
KW   Crohn's disease; ulcerative colitis.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 103
FT                   /label= OTHER
FT                   /note= "Other- STOP"
XX
PN   WO200131037-A2.
XX
PD   03-MAY-2001.
XX
PF   26-OCT-2000; 2000WO-IL00688.
XX
PR   27-OCT-1999;   99IL-0132611.
XX
PA   (YEDA ) YEDA RES & DEV CO LTD.
XX
PI   Ben-Nun A,  Kerlero De Rosbo N,  Sappler GP;
XX
DR   WPI; 2001-300515/31.
DR   N-PSDB; AAS04827.
XX
PT   Novel synthetic human target autoantigen gene useful for treating
PT   autoimmune diseases such as multiple sclerosis, insulin-dependent
PT   diabetes mellitus, rheumatoid arthritis, myasthenia gravis, and uveitis
PT   -
XX
PS   Claim 173; Fig 51; 182pp; English.
XX
CC   The sequence represents shGLP/RA (synthetic human chondrocyte
CC   glycoprotein 69/cartilage link protein) containing immunogenic epitope
CC   clusters (IEC) from chondrocyte glycoprotein 69 and cartilage link
CC   protein. The synthetic human target autoantigen genes of the invention
CC   comprise sequences coding for at least 2 IECs of autoantigen(s) related
CC   to a specific autoimmune disease. The synthetic human target autoantigen
CC   genes are useful for treating autoimmune diseases such as multiple
CC   sclerosis, insulin-dependent diabetes mellitus, rheumatoid arthritis,
CC   myasthenia gravis, uveitis, autoimmune hepatitis, thyroiditis, insulitis,
CC   orchitis, idiopathic thrombocytopenic purpura, and inflammatory diseases
CC   (Crohn's disease, ulcerative colitis). The synthetic human target
CC   autoantigen genes are also useful for diagnosis and/or monitoring the
CC   progression of the autoimmune disease.

XX
SQ   Sequence   105 AA;

  Query Match             100.0%;  Score 65;  DB 22;  Length 105;
  Best Local Similarity   100.0%;  Pred. No. 0.00029;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VGYDDQESVKSKV 13
              |||||||||||||
Db         45 VGYDDQESVKSKV 57


RESULT 5
AAU02078
ID   AAU02078 standard; Protein; 220 AA.
XX
AC   AAU02078;
XX
DT   07-SEP-2001  (first entry)
XX
DE   Synthetic human multi-target autoantigen Y-RAPc.
XX
KW   Y-RAPc; immunogenic epitope cluster; IEC;
KW   autoantigen; shCollagen/RA; shAggrecan/RA; shGLP/RA;
KW   autoimmune disease; multiple sclerosis; rheumatoid arthritis;
KW   insulin-dependent diabetes mellitus; myasthenia gravis; uveitis;
KW   autoimmune hepatitis; thyroiditis; insulitis; orchitis;
KW   idiopathic thrombocytopenic purpura; inflammatory disease;
KW   Crohn's disease; ulcerative colitis; shMultiTAG.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 218
FT                   /label= OTHER
FT                   /note= "Other- STOP"
XX
PN   WO200131037-A2.
XX
PD   03-MAY-2001.
XX
PF   26-OCT-2000; 2000WO-IL00688.
XX
PR   27-OCT-1999;   99IL-0132611.
XX
PA   (YEDA ) YEDA RES & DEV CO LTD.
XX
PI   Ben-Nun A,  Kerlero De Rosbo N,  Sappler GP;
XX
DR   WPI; 2001-300515/31.
DR   N-PSDB; AAS04839.
XX
PT   Novel synthetic human target autoantigen gene useful for treating
PT   autoimmune diseases such as multiple sclerosis, insulin-dependent
PT   diabetes mellitus, rheumatoid arthritis, myasthenia gravis, and uveitis
PT   -
XX
PS   Claim 185; Fig 59; 182pp; English.
XX
CC   The sequence represents synthetic human multi-target autoantigen
CC   (shMultiTAG) Y-RAPa consisting of regions encoding preferred epitopes
CC   from shCollagen/RA, shAggrecan/RA and shGLP/RA. The synthetic human
CC   target autoantigen genes of the invention comprise sequences coding for
CC   at least 2 IECs of autoantigen(s) related to a specific autoimmune
CC   disease. The synthetic human target autoantigen genes are useful for
CC   treating autoimmune diseases such as multiple sclerosis, insulin-
CC   dependent diabetes mellitus, rheumatoid arthritis, myasthenia gravis,
CC   uveitis, autoimmune hepatitis, thyroiditis, insulitis, orchitis,
CC   idiopathic thrombocytopenic purpura, and inflammatory diseases (Crohn's
CC   disease, ulcerative colitis). The synthetic human target autoantigen
CC   genes are also useful for diagnosis and/or monitoring the progression of
CC   the autoimmune disease.
```

```
XX
SQ   Sequence   220 AA;

  Query Match             100.0%;  Score 65;  DB 22;  Length 220;
  Best Local Similarity   100.0%;  Pred. No. 0.00068;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VGYDDQESVKSKV 13
              |||||||||||||
Db        201 VGYDDQESVKSKV 213


RESULT 6
AAU02079
ID   AAU02079 standard; Protein; 281 AA.
XX
AC   AAU02079;
XX
DT   07-SEP-2001  (first entry)
XX
DE   Synthetic multi-target autoantigen Y-RAPd.
XX
KW   Y-RAPd; immunogenic epitope cluster; IEC;
KW   autoantigen; shCollagen/RA; shAggrecan/RA; shGLP/RA; shHSP/RA;
KW   autoimmune disease; multiple sclerosis; rheumatoid arthritis;
KW   insulin-dependent diabetes mellitus; myasthenia gravis; uveitis;
KW   autoimmune hepatitis; thyroiditis; insulitis; orchitis;
KW   idiopathic thrombocytopenic purpura; inflammatory disease;
KW   Crohn's disease; ulcerative colitis; shMultiTAG.
XX
OS   Synthetic.
OS   Homo sapiens.
OS   Escherichia coli.
OS   Mycobacterium tuberculosis.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 279
FT                   /label= OTHER
FT                   /note= "Other- STOP"
XX
PN   WO200131037-A2.
XX
PD   03-MAY-2001.
XX
PF   26-OCT-2000; 2000WO-IL00688.
XX
PR   27-OCT-1999;   99IL-0132611.
XX
PA   (YEDA ) YEDA RES & DEV CO LTD.
XX
PI   Ben-Nun A.  Kerlero De Rosbo N,  Sappler GP;
XX
DR   WPI; 2001-300515/31.
DR   N-PSDB; AAS04840.
XX
PT   Novel synthetic human target autoantigen gene useful for treating
PT   autoimmune diseases such as multiple sclerosis, insulin-dependent
PT   diabetes mellitus, rheumatoid arthritis, myasthenia gravis, and uveitis
PT   -
XX
PS   Claim 188; Fig 61; 182pp; English.
XX
CC   The sequence represents synthetic human multi-target autoantigen
CC   (shMultiTAG) Y-RAPd consisting of preferred epitopes encoded by
CC   shCollagen/RA, shAggrecan/RA, shGLP/RA and shHSP/RA. The synthetic human
CC   target autoantigen genes of the invention comprise sequences coding for
CC   at least 2 IECs of autoantigen(s) related to a specific autoimmune
CC   disease. The synthetic human target autoantigen genes are useful for
CC   treating autoimmune diseases such as multiple sclerosis, insulin-
CC   dependent diabetes mellitus, rheumatoid arthritis, myasthenia gravis,
CC   uveitis, autoimmune hepatitis, thyroiditis, insulitis, orchitis,
CC   idiopathic thrombocytopenic purpura, and inflammatory diseases (Crohn's
CC   disease, ulcerative colitis). The synthetic human target autoantigen
CC   genes are also useful for diagnosis and/or monitoring the progression of
CC   the autoimmune disease.
XX
SQ   Sequence   281 AA;

  Query Match             100.0%;  Score 65;  DB 22;  Length 281;
  Best Local Similarity   100.0%;  Pred. No. 0.0009;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VGYDDQESVKSKV 13
              |||||||||||||
Db        201 VGYDDQESVKSKV 213


RESULT 7
AAP81342
ID   AAP81342 standard; protein; 321 AA.
XX
AC   AAP81342;
XX
DT   25-MAR-2003  (updated)
DT   10-MAR-2003  (updated)
DT   19-OCT-1990  (first entry)
XX
DE   Polypeptide involved in protective mechanisms.
XX
KW   Immune response; cell growth.
XX
OS   Unidentified.
XX
PN   JP63023898-A.
XX
PD   30-JAN-1988.
XX
PF   16-JUL-1986;   86JP-0167518.
XX
PR   16-JUL-1986;   86JP-0167518.
XX
PA   (DAIN ) DAINIPPON PHARM KK.
XX
DR   WPI; 1988-068419/10.
DR   P-PSDB; AAP81342.
XX
PT   New polypeptide and DNA encoding it - related to protective
PT   mechanisms such as immune response etc.
XX
PS   Disclosure; Page ?; ?pp; Japanese.
XX
CC   This polypeptide is involved in protective mechanisms such as immune
CC   response, cell growth and activation of protective functions.
CC   (Updated on 10-MAR-2003 to add missing OS field.)
CC   (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence   321 AA;

  Query Match             100.0%;  Score 65;  DB 9;  Length 321;
  Best Local Similarity   100.0%;  Pred. No. 0.0011;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VGYDDQESVKSKV 13
              |||||||||||||
Db        264 VGYDDQESVKSKV 276


RESULT 8
AAU02076
ID   AAU02076 standard; Protein; 350 AA.
XX
AC   AAU02076;
XX
DT   07-SEP-2001  (first entry)
XX
DE   Synthetic human multi-target autoantigen Y-RAPa.
```

```
XX
KW   Y-RAPa; immunogenic epitope cluster; IEC;
KW   autoantigen; shCollagen/RA; shAggrecan/RA; shGLP/RA;
KW   autoimmune disease; multiple sclerosis; rheumatoid arthritis;
KW   insulin-dependent diabetes mellitus; myasthenia gravis; uveitis;
KW   autoimmune hepatitis; thyroiditis; insulitis; orchitis;
KW   idiopathic thrombocytopenic purpura; inflammatory disease;
KW   Crohn's disease; ulcerative colitis; shMultiTAG.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 348
FT                   /label= OTHER
FT                   /note= "Other- STOP"
XX
PN   WO200131037-A2.
XX
PD   03-MAY-2001.
XX
PF   26-OCT-2000; 2000WO-IL00688.
XX
PR   27-OCT-1999;   99IL-0132611.
XX
PA   (YEDA ) YEDA RES & DEV CO LTD.
XX
PI   Ben-Nun A,  Kerlero De Rosbo N,  Sappler GP;
XX
DR   WPI; 2001-300515/31.
DR   N-PSDB; AAS04837.
XX
PT   Novel synthetic human target autoantigen gene useful for treating
PT   autoimmune diseases such as multiple sclerosis, insulin-dependent
PT   diabetes mellitus, rheumatoid arthritis, myasthenia gravis, and uveitis
PT   -
XX
PS   Claim 179; Fig 55; 182pp; English.
XX
CC   The sequence represents synthetic human multi-target autoantigen
CC   (shMultiTAG) Y-RAPa consisting of regions encoding shCollagen/RA,
CC   shAggrecan/RA and shGLP/RA. The synthetic human target autoantigen genes
CC   of the invention comprise sequences coding for at least 2 IECs of
CC   autoantigen(s) related to a specific autoimmune disease. The synthetic
CC   human target autoantigen genes are useful for treating autoimmune
CC   diseases such as multiple sclerosis, insulin-dependent diabetes
CC   mellitus, rheumatoid arthritis, myasthenia gravis, uveitis, autoimmune
CC   hepatitis, thyroiditis, insulitis, orchitis, idiopathic thrombocytopenic
CC   purpura, and inflammatory diseases (Crohn's disease, ulcerative colitis).
CC   The synthetic human target autoantigen genes are also useful for
CC   diagnosis and/or monitoring the progression of the autoimmune disease.
XX
SQ   Sequence   350 AA;

  Query Match           100.0%;  Score 65;  DB 22;  Length 350;
  Best Local Similarity  100.0%;  Pred. No. 0.0012;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VGYDDQESVKSKV 13
              |||||||||||||
Db        290 VGYDDQESVKSKV 302


RESULT 9
ABG96297
ID   ABG96297 standard; Protein; 383 AA.
XX
AC   ABG96297;
XX
DT   11-DEC-2002  (first entry)
XX
DE   Human ovarian cancer marker OV6.
XX
KW   Human; ovarian cancer; marker; cancer; familial history; brain disorder;
KW   central nervous system disorder; bacterial meningitis; viral meningitis;
KW   Alzheimer's disease; Parkinson's disease; cerebral oedema; hydrocephalus;
KW   brain herniation; inflammation; encephalitis; testicular disorder;
KW   nontuberculous granulomatous orchitis; connective tissue disorder;
KW   heart disorder; ischaemic heart disease; atherosclerosis; neoplasm;
KW   histological type; carcinogenic; ovarian cancer marker.
XX
OS   Homo sapiens.
XX
PN   WO200271928-A2.
XX
PD   19-SEP-2002.
XX
PF   14-MAR-2002; 2002WO-US07826.
XX
PR   14-MAR-2001; 2001US-276025P.
PR   14-MAR-2001; 2001US-276026P.
PR   10-AUG-2001; 2001US-311732P.
PR   19-SEP-2001; 2001US-323580P.
PR   26-SEP-2001; 2001US-324967P.
PR   26-SEP-2001; 2001US-325102P.
PR   26-SEP-2001; 2001US-325149P.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   Monahan JE,  Gannavarapu M,  Boersch S,  Kamatkar S,  Kovatis SG;
PI   Meyers RE,  Morrisey MP,  Olandt PJ,  Sen A,  Vieby PG,  Mills GB;
PI   Bast RC,  Lu K,  Schmandt RE,  Zhao X,  Glatt K;
XX
DR   WPI; 2002-723277/78.
DR   N-PSDB; ABS76389.
XX
PT   Assessing whether a patient is afflicted with ovarian cancer, useful in
PT   assessing the stage or progression of the disease, comprises comparing
PT   the expression level of a cancer marker in a sample from a patient and
PT   from a non cancer patient -
XX
PS   Disclosure; Page 164-165; 481pp; English.
XX
CC   The present invention relates to a new method for assessing whether a
CC   patient is afflicted with ovarian cancer. The method involves comparing
CC   the expression level of a marker in a patient sample and the normal level
CC   of expression of the marker in a control non-ovarian cancer sample, where
CC   the marker is selected from 363 cancer markers described in the
CC   specification. The method of the invention is useful in diagnosing or
CC   characterising cancer, in detecting the presence of cancer as early as
CC   possible, and the recurrence of ovarian cancer. The method may also be of
CC   particular use with patients having an enhanced risk of developing
CC   ovarian cancer (e.g. patients having a familial history of ovarian
CC   cancer). The cancer markers may be used in the management and treatment
CC   of e.g. brain and central nervous system disorders (e.g. bacterial and
CC   viral meningitis, Alzheimer's disease or Parkinson's disease), brain
CC   disorders (e.g. cerebral oedema, hydrocephalus or brain herniations),
CC   inflammations (e.g. bacterial or viral meningitis or encephalitis),
CC   testicular disorders (e.g. nontuberculous granulomatous orchitis),
CC   connective tissue disorders, or heart disorders (e.g. ischaemic heart
CC   disease or atherosclerosis). The compositions and methods may also be
CC   used in assessing the histological type of neoplasm associated with
CC   ovarian cancer, monitoring the progression of ovarian cancer,
CC   determining whether ovarian cancer has metastasized or is likely to
CC   metastasize, selecting a composition for inhibiting ovarian cancer,
CC   assessing the ovarian carcinogenic potential of a compound, or
CC   inhibiting ovarian cancer or at risk of developing ovarian cancer. The
CC   present amino acid sequence represents one of the ovarian cancer markers
CC   described in the invention.
XX
SQ   Sequence   383 AA;

  Query Match           100.0%;  Score 65;  DB 23;  Length 383;
  Best Local Similarity  100.0%;  Pred. No. 0.0013;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 VGYDDQESVKSKV 13
              ||:||||||||||
Db        326 VGYDDQESVKSKV 338


RESULT 10
ABU56651
ID   ABU56651 standard; Protein; 383 AA.
XX
AC   ABU56651;
XX
DT   02-APR-2003  (first entry)
XX
DE   Lung cancer-associated polypeptide #244.
XX
KW   Lung cancer-associated polypeptide; cytostatic; emphysema;
KW   antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
KW   small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
KW   chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
KW   interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
XX
OS   Unidentified.
XX
PN   WO200286443-A2.
XX
PD   31-OCT-2002.
XX
PF   18-APR-2002; 2002WO-US12476.
XX
PR   18-APR-2001; 2001US-284770P.
PR   10-MAY-2001; 2001US-290492P.
PR   09-NOV-2001; 2001US-339245P.
PR   13-NOV-2001; 2001US-350666P.
PR   29-NOV-2001; 2001US-334370P.
PR   12-APR-2002; 2002US-372246P.
XX
PA   (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI   Aziz N,  Murray R;
XX
DR   WPI; 2003-093161/08.
DR   N-PSDB; ABX76380.
XX
PT   Detecting a lung cancer-associated transcript in a cell from a patient
PT   for treating lung cancer, by contacting a biological sample from the
PT   patient with a polynucleotide that exhibits increased or decreased
PT   expression in lung cancer
XX
PS   Claim 27; Page 377; 453pp; English.
XX
CC   The invention relates to a method for detecting a lung cancer-associated
CC   transcript in a cell from a patient, comprising contacting a biological
CC   sample from the patient with a polynucleotide that selectively hybridises
CC   to a sequence that is at least 80 % identical to a gene that exhibits
CC   increased or decreased expression in lung cancer samples. Lung
CC   cancer-associated polynucleotides and polypeptides are used for
CC   identifying a compound that modulates a lung cancer-associated
CC   polypeptide, for inhibiting proliferation of a lung cancer-associated
CC   cell to treat lung cancer in a patient and for treating a mammal having
CC   lung cancer by administering a modulatory compound identified. The
CC   methods are useful for treating lung cancer, such as small cell lung
CC   cancer, non-small cell lung cancer or other benign or precancerous
CC   lesions, e.g. atelectasis, emphysema, bronchitis, chronic obstructive
CC   pulmonary disease, fibrosis, hypersensitivity pneumonitis, interstitial
CC   pulmonary fibrosis, asthma and bronchiectasis. The genes, polynucleotides
CC   and polypeptides are useful for diagnostic purposes and as targets for
CC   screening for therapeutic compounds that modulate lung cancer, such as
CC   antibodies. Sequences ABU56408-ABU56745 represent lung cancer-associated
CC   polypeptides of the invention.
XX
SQ   Sequence   383 AA;

  Query Match           100.0%;  Score 65;  DB 24;  Length 383;
  Best Local Similarity   100.0%;  Pred. No. 0.0013;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VGYDDQESVKSKV 13
              |||||||||||||
Db        326 VGYDDQESVKSKV 338


RESULT 11
AAU02077
ID   AAU02077 standard; Protein; 459 AA.
XX
AC   AAU02077;
XX
DT   07-SEP-2001  (first entry)
XX
DE   Synthetic multi-target autoantigen Y-RAPb.
XX
KW   Y-RAPa; immunogenic epitope cluster; IEC;
KW   autoantigen; shCollagen/RA; shAggrecan/RA; shGLP/RA; shHSP/RA;
KW   autoimmune disease; multiple sclerosis; rheumatoid arthritis;
KW   insulin-dependent diabetes mellitus; myasthenia gravis; uveitis;
KW   autoimmune hepatitis; thyroiditis; insulitis; orchitis;
KW   idiopathic thrombocytopenic purpura; inflammatory disease;
KW   Crohn's disease; ulcerative colitis; shMultiTAG.
XX
OS   Synthetic.
OS   Homo sapiens.
OS   Escherichia coli.
OS   Mycobacterium tuberculosis.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 455
FT                   /label= OTHER
FT                   /note= "Other= STOP"
FT   Misc-difference 457
FT                   /label= OTHER
FT                   /note= "Other= STOP"
XX
PN   WO200131037-A2.
XX
PD   03-MAY-2001.
XX
PF   26-OCT-2000; 2000WO-IL00688.
XX
PR   27-OCT-1999;  99IL-0132611.
XX
PA   (YEDA ) YEDA RES & DEV CO LTD.
XX
PI   Ben-Nun A,  Kerlero De Rosbo N,  Sappler GP;
XX
DR   WPI; 2001-300515/31.
DR   N-PSDB; AAS04838.
XX
PT   Novel synthetic human target autoantigen gene useful for treating
PT   autoimmune diseases such as multiple sclerosis, insulin-dependent
PT   diabetes mellitus, rheumatoid arthritis, myasthenia gravis, and uveitis
PT
XX
PS   Claim 182; Fig 57; 182pp; English.
XX
CC   The sequence represents synthetic human multi-target autoantigen
CC   (shMultiTAG) Y-RAPa consisting of regions encoding shCollagen/RA,
CC   shAggrecan/RA, shGLP/RA and shHSP/RA. The synthetic human target
CC   autoantigen genes of the invention comprise sequences coding for at least
CC   2 IECs of autoantigen(s) related to a specific autoimmune disease. The
CC   synthetic human target autoantigen genes are useful for treating
CC   autoimmune diseases such as multiple sclerosis, insulin-dependent
CC   diabetes mellitus, rheumatoid arthritis, myasthenia gravis, uveitis,
CC   autoimmune hepatitis, thyroiditis, insulitis, orchitis, idiopathic
CC   thrombocytopenic purpura, and inflammatory diseases (Crohn's disease,
CC   ulcerative colitis). The synthetic human target autoantigen genes are
CC   also useful for diagnosis and/or monitoring the progression of the
```

```
CC   autoimmune disease.
XX
SQ   Sequence   459 AA;

  Query Match           100.0%;  Score 65;  DB 22;  Length 459;
  Best Local Similarity  100.0%;  Pred. No. 0.0016;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VGYDDQESVKSKV 13
              |||||||||||||
Db        290 VGYDDQESVKSKV 302


RESULT 12
AAW45150
ID   AAW45150 standard; peptide; 16 AA.
XX
AC   AAW45150;
XX
DT   28-APR-1998  (first entry)
XX
DE   Human cartilage glycoprotein 39 derived peptide #51.
XX
KW   Articular cartilage; immunosuppressive therapy; antigen; autoantigen;
KW   immunological tolerance; T-cell; human cartilage glycoprotein 39;
KW   HC gp-39; rheumatoid arthritis; epitope.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9740068-A1.
XX
PD   30-OCT-1997.
XX
PF   22-APR-1997;   97WO-EP02051.
XX
PR   24-APR-1996;   96EP-0201106.
XX
PA   (ALKO ) AKZO NOBEL NV.
XX
PI   Boots AMH,  Verheijden GFM;
XX
DR   WPI; 1997-535775/49.
XX
PT   Peptide suitable for use in antigen specific immunosuppressive
PT   therapy - resembles or mimics epitope present on HC gp-39, so
PT   inducing systemic immunological tolerance to rheumatoid arthritis
PT   auto:antigen
XX
PS   Disclosure; Page 17; 82pp; English.
XX
CC   The present sequence represents a peptide which resembles or mimics an
CC   epitope present on human cartilage glycoprotein 39 (HC gp-39), an
CC   autoantigen in rheumatoid arthritis. The invention relates to peptides
CC   consisting of 16-55 amino acid residues comprising at least one of the
CC   following 19 sequences: LVCYYTSWS; FLCTHIIYS; IIYSFANIS; LRTILSVGC;
CC   FIKSVPPFL; FDGLDLAWL; LYPGRRDKQ; YDIAKISQH; LDFISIMTY; FISIMTYDF;
CC   FRGQEDASP; YAVGYMLRL; MLRLGAPAS; LAYYEICDF; LRGATVHRI; YLKDRQLAG;
CC   LAGAMVWAL; VWALDLDDF; or LDLDDFQGS. They can be used medically in
CC   antigen specific immunosuppressive therapy, particularly the treatment
CC   of T-cell mediated destruction of articular cartilage in autoimmune
CC   diseases (e.g. rheumatoid arthritis). They can also be used to detect
CC   activated autoreactive T cells in an individual. The peptides have a
CC   specific effect on the autoreactive T cells, thus leaving the other
CC   components of the immune system intact, unlike the non-specific
CC   suppressive effect of immunosuppressive drugs, and do not cause toxic
CC   side effects. The peptides are predominantly recognised by autoreactive
CC   T cells from rheumatoid arthritis patients, but rarely by those from
CC   healthy donors.
XX
SQ   Sequence   16 AA;

  Query Match            93.8%;  Score 61;  DB 18;  Length 16;
  Best Local Similarity  100.0%;  Pred. No. 0.00018;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VGYDDQESVKSK 12
              ||||||||||||
Db          5 VGYDDQESVKSK 16


RESULT 13
AAB03442
ID   AAB03442 standard; Protein; 383 AA.
XX
AC   AAB03442;
XX
DT   03-JAN-2001  (first entry)
XX
DE   Gp38k protein sequence.
XX
KW   Gp38k; chemoattractant; cell migration; wound healing; angiogenesis;
KW   cancer; vascular trauma; vascular disease; atherosclerosis; restenosis;
XX
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..24
FT                   /label= signal_peptide
FT   Region          30..37
FT                   /note= "peptide antibody"
FT   Modified-site   59..61
FT                   /label= glycosylation_site
FT   Binding-site    68..75
FT                   /label= leucine_zipper
FT   Binding-site    82..89
FT                   /label= leucine_zipper
FT   Active-site     131..136
FT                   /label= chitanase_active_site
FT   Binding-site    143..146
FT                   /label= heparin_binding_site
FT   Binding-site    147..154
FT                   /label= hyaluronic_acid_binding_site
FT   Binding-site    262..270
FT                   /label= hyaluronic_acid_binding_site
FT   Binding-site    278..281
FT                   /label= glycosaminoglycan_binding_site
FT   Region          354..357
FT                   /note= "acidic region"
FT   Binding-site    368..376
FT                   /label= hyaluronic_acid_binding_site
XX
PN   WO200034469-A1.
XX
PD   15-JUN-2000.
XX
PF   10-DEC-1999;   99WO-US29262.
XX
PR   11-DEC-1998;   98US-0111856.
XX
PA   (UYNY ) UNIV NEW YORK STATE RES FOUND.
XX
PI   Millis AJT;
XX
DR   WPI; 2000-431300/37.
XX
PT   Clusterin and gp38K-related peptide capable of altering cell migration
PT   useful for treating atherosclerosis, cancer and stenosis following
PT   vascular trauma or disease -
XX
PS   Disclosure; Fig 2; 43pp; English.
XX
CC   The present sequence is the protein sequence of gp38k. Gp38k, a
CC   chemoattractant, is essential for the migration of vascular smooth muscle
CC   cells (VSMC). The gene and protein can, therefore, be used to promote
CC   wound healing, angiogenesis and vasculogenesis, in the treatment of
```

```
CC   stenosis following vascular trauma or disease and to treat
CC   atherosclerosis, and antisense sequences can be used to treat cancer, as
CC   angiogenesis is vital for tumour survival.
XX
SQ   Sequence   383 AA;

  Query Match             89.2%;  Score 55;  DB 21;  Length 383;
  Best Local Similarity   91.7%;  Pred. No. 0.025;
  Matches   11;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VGYDDQESVKSK 12
              |||||||||:|
Db        326 VGYDDQESVKNK 337


RESULT 14
AAW45151
ID   AAW45151 standard; peptide; 16 AA.
XX
AC   AAW45151;
XX
DT   28-APR-1998  (first entry)
XX
DE   Human cartilage glycoprotein 39 derived peptide #52.
XX
KW   Articular cartilage; immunosuppressive therapy; antigen; autoantigen;
KW   immunological tolerance; T-cell; human cartilage glycoprotein 39;
KW   HC gp-39; rheumatoid arthritis; epitope.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9740068-A1.
XX
PD   30-OCT-1997.
XX
PF   22-APR-1997;   97WO-EP02051.
XX
PR   24-APR-1996;   96EP-0201106.
XX
PA   (ALKU ) AKZO NOBEL NV.
XX
PI   Boots AMH,  Verheijden GFM;
XX
DR   WPI; 1997-535775/49.
XX
PT   Peptide suitable for use in antigen specific immunosuppressive
PT   therapy - resembles or mimics epitope present on HC gp-39, so
PT   inducing systemic immunological tolerance to rheumatoid arthritis
PT   auto:antigen
XX
PS   Disclosure; Page 17; 82pp; English.
XX
CC   The present sequence represents a peptide which resembles or mimics an
CC   epitope present on human cartilage glycoprotein 39 (HC gp-39), an
CC   autoantigen in rheumatoid arthritis. The invention relates to peptides
CC   consisting of 16-55 amino acid residues comprising at least one of the
CC   following 19 sequences: LVCYYTSWS; FLCTHIIYS; IIYSFANIS; LKILLSVGG;
CC   FIKSVPPFL; FDGLDLAWL; LYPGRRDKQ; YDIAKISQH; LDFISIMTY; FISIMTYDF;
CC   FRGQEDASP; YAVGYMLRL; MLRLGAPAS; LAYYEICDP; LRGATVHRT; YLKDRQLAG;
CC   LAGAMVWAL; VWALDLDDF; or LDLDDFQGS. They can be used medically in
CC   antigen specific immunosuppressive therapy, particularly the treatment
CC   of T-cell mediated destruction of articular cartilage in autoimmune
CC   diseases (e.g. rheumatoid arthritis). They can also be used to detect
CC   activated autoreactive T cells in an individual. The peptides have a
CC   specific effect on the autoreactive T cells, thus leaving the other
CC   components of the immune system intact, unlike the non-specific
CC   suppressive effect of immunosuppressive drugs, and do not cause toxic
CC   side effects. The peptides are predominantly recognised by autoreactive
CC   T cells from rheumatoid arthritis patients, but rarely by those from
CC   healthy donors.
XX
SQ   Sequence   16 AA;

  Query Match             84.6%;  Score 55;  DB 18;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 0.0023;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 YDDQESVKSKV 13
              |||:||||||
Db          1 YDDQESVKSKV 11


RESULT 15
ABB55896
ID   ABB55896 standard; Peptide; 14 AA.
XX
AC   ABB55896;
XX
DT   15-FEB-2002  (first entry)
XX
DE   Vascular dementia-associated protein isoform (VPI) 96.
XX
KW   Vascular Dementia; VD; VD-associated protein isoform; VPI; screening;
KW   diagnosis; prognosis; gene therapy.
XX
OS   Homo sapiens.
XX
PN   WO200169261-A2.
XX
PD   20-SEP-2001.
XX
PF   14-MAR-2001; 2001WO-GB01106.
XX
PR   15-MAR-2000; 2000GB-0006285.
PR   24-NOV-2000; 2000GB-0028734.
PR   28-NOV-2000; 2000US-0724397.
XX
PA   (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
XX
PI   Herath HMAC,  Parekh RB,  Rohlff C;
XX
DR   WPI; 2001-557937/62.
XX
PT   Screening, diagnosis or prognosis of vascular dementia (VD), useful for
PT   determining stage of VD and monitoring the effect of VD therapy,
PT   comprises analysing body fluid by 2-dimensional electrophoresis for
PT   features correlated with VD -
XX
PS   Claim 6; Page 31; 151pp; English.
XX
CC   The invention relates to screening, diagnosis or prognosis of Vascular
CC   Dementia (VD) in a subject comprising analysing body fluid from the
CC   subject by 2-dimensional (2-D) electrophoresis to generate a 2-D array of
CC   features containing at least one chosen feature whose relative abundance
CC   correlates with the presence, absence, stage or severity of VD or
CC   predicts the onset or course of VD, especially detecting in a sample of
CC   cerebrospinal fluid (CSF) from the subject one of 223 VD-associated
CC   protein isoforms (VPIs) (ABB55801-ABB56295) as fully defined in the
CC   specification. Detecting VD-associated features and VPI is useful for the
CC   screening, diagnosis or prognosis of VD, for determining the stage or
CC   severity of VD, for identifying a subject at risk of VD or for
CC   monitoring the effect of therapy administered to a subject having VD.
CC   Nucleic acids encoding a VPI or inhibiting the function of a VPI are
CC   useful for the treatment of VD and for gene therapy.
XX
SQ   Sequence   14 AA;

  Query Match             80.0%;  Score 52;  DB 22;  Length 14;
  Best Local Similarity   100.0%;  Pred. No. 0.007;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VGYDDQESVK 10
              ||||||||||
Db          5 VGYDDQESVK 14
```

Search completed: September 24, 2003, 13:41:23
Job time : 53.2711 secs

GenCore version 5.1.6

OM protein - protein search, using sw mode.

Run on: September 24, 2003, 13:35:23 ; Search time 15.7753 Seconds
(without alignments)
34.667 Million cell updates/sec

Title: US-09-744-282-7
Perfect score: 65
Sequence: 1 VGYDDQESVKSKV 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	65	100.0	13	1	US-08-619-645-7	Sequence 7, Appli
2	65	100.0	13	2	US-08-634-493-7	Sequence 7, Appli
3	61	93.8	16	3	US-09-171-705-51	Sequence 51, Appl
4	58	89.2	383	4	US-09-459-749D-17	Sequence 17, Appl
5	55	84.6	16	3	US-09-171-705-52	Sequence 52, Appl
6	47	72.3	739	4	US-09-503-922-1	Sequence 1, Appli
7	46	70.8	9	1	US-08-619-645-3	Sequence 3, Appli
8	46	70.8	9	2	US-08-634-493-3	Sequence 3, Appli
9	46	70.8	373	3	US-09-039-198A-14	Sequence 14, Appl
10	46	70.8	373	3	US-09-039-198A-15	Sequence 15, Appl
11	46	70.8	373	4	US-08-877-599-14	Sequence 14, Appl
12	46	70.8	373	4	US-08-877-599-15	Sequence 15, Appl
13	46	70.8	373	4	US-09-267-574-14	Sequence 14, Appl
14	46	70.8	373	4	US-09-267-574-15	Sequence 15, Appl
15	46	70.8	387	2	US-08-486-839-6	Sequence 6, Appli
16	46	70.8	387	3	US-09-151-011-6	Sequence 6, Appli
17	46	70.8	387	4	US-09-343-623-6	Sequence 6, Appli
18	46	70.8	466	2	US-08-486-839-4	Sequence 4, Appli
19	46	70.8	466	3	US-09-151-011-4	Sequence 4, Appli
20	46	70.8	466	3	US-09-039-198A-2	Sequence 2, Appli
21	46	70.8	466	3	US-09-039-198A-4	Sequence 4, Appli
22	46	70.8	466	4	US-09-343-623-4	Sequence 4, Appli
23	46	70.8	466	4	US-08-877-599-2	Sequence 2, Appli
24	46	70.8	466	4	US-08-877-599-4	Sequence 4, Appli
25	46	70.8	466	4	US-09-267-574-2	Sequence 2, Appli
26	46	70.8	466	4	US-09-267-574-4	Sequence 4, Appli
27	44	67.7	385	2	US-08-694-915-2	Sequence 2, Appli
28	44	67.7	416	2	US-08-694-915-4	Sequence 4, Appli
29	44	67.7	423	4	US-08-850-348A-2	Sequence 2, Appli
30	43	66.2	371	2	US-08-591-629-2	Sequence 2, Appli
31	40	61.5	377	2	US-08-591-629-8	Sequence 8, Appli
32	38	58.5	490	4	US-09-292-225-41	Sequence 41, Appl
33	38	58.5	509	4	US-09-292-225-35	Sequence 35, Appl
34	38	58.5	509	4	US-09-292-225-38	Sequence 38, Appl
35	38	58.5	536	4	US-09-292-225-21	Sequence 21, Appl
36	38	58.5	555	4	US-09-292-225-15	Sequence 15, Appl
37	38	58.5	555	4	US-09-292-225-18	Sequence 18, Appl
38	37	56.9	282	4	US-09-107-532A-4057	Sequence 4057, Ap
39	37	56.9	400	2	US-08-713-298B-2	Sequence 2, Appli
40	37	56.9	400	2	US-08-870-180B-2	Sequence 2, Appli
41	37	56.9	400	3	US-08-814-052-4	Sequence 4, Appli
42	37	56.9	400	3	US-08-812-629-4	Sequence 4, Appli
43	37	56.9	400	3	US-09-226-529-2	Sequence 2, Appli
44	37	56.9	462	2	US-08-870-180B-13	Sequence 13, Appl
45	37	56.9	462	3	US-09-226-529-13	Sequence 13, Appl

ALIGNMENTS

RESULT 1
US-08-619-645-7
; Sequence 7, Application US/08619645
; Patent No. 5736507
; GENERAL INFORMATION:
; APPLICANT: Boots, Anna M.H.
; APPLICANT: Verheijden, Gijsbertus F.M.
; TITLE OF INVENTION: No. 5736507el peptides derived from
; TITLE OF INVENTION: autoantigen for use in immunotherapy of autoimmune
; TITLE OF INVENTION: disease
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Akzo No. 5736507el Patent Department
; STREET: 1300 Piccard Drive, Suite 206
; CITY: Rockville
; STATE: Maryland
; COUNTRY: U.S.A.
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25(EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/619,645
; FILING DATE: 25-MAR-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Gormley, Mary E.
; REGISTRATION NUMBER: 34,409
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 947-4433
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-619-645-7

Query Match 100.0%; Score 65; DB 1; Length 13;
Best Local Similarity 100.0%; Pred. No. 1.7e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VGYDDQESVKSKV 13
||||||||||||||
Db 1 VGYDDQESVKSKV 13

```
RESULT 2
US-08-634-493-7
; Sequence 7, Application US/08634493
; Patent No. 5843449
;  GENERAL INFORMATION:
;    APPLICANT:  A.M.H. Boots
;    APPLICANT:  G.F.M. Verheijden
;    APPLICANT:  E.S. Bos
;    TITLE OF INVENTION:  No. 5843449el Peptides derived from autoantigen for use
;    TITLE OF INVENTION:  in Immunotherapy of Autoimmune Diseases
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Akzo No. 5843449el Patent Department
;      STREET:  1300 Piccard Drive, Suite 206
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/634,493
;      FILING DATE:  18-APR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/619,645
;      FILING DATE:  25-MAR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/EP95/04201
;      FILING DATE:  25-OCT-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  NL 9420381287
;      FILING DATE:  27-OCT-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  NL 952008860
;      FILING DATE:  07-APR-1995
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 258-5200
;      TELEFAX:  (301) 977-0847
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  13 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-634-493-7

  Query Match             100.0%;  Score 65;  DB 2;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-05;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VGYDDQESVKSKV 13
              |||||||||||||
Db          1 VGYDDQESVKSKV 13


RESULT 3
US-09-171-705-51
; Sequence 51, Application US/09171705
; Patent No. 6184204
; GENERAL INFORMATION:
;  APPLICANT: BOOTS, ANNA M.H.
;  APPLICANT: VERHEIJDEN, GILBERTUS F.M.
;  TITLE OF INVENTION: NOVEL PEPTIDES SUITABLE FOR USE IN ANTIGEN SPECIFIC
;  TITLE OF INVENTION: IMMUNOSUPPRESSIVE THERAPY
;  FILE REFERENCE: O/96198 US
;  CURRENT APPLICATION NUMBER: US/09/171,705
;  CURRENT FILING DATE: 1999-02-09
;  NUMBER OF SEQ ID NOS: 78
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 51
;   LENGTH: 16
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: DERIVED FROM
;   OTHER INFORMATION: SEQUENCE OF HUMAN CARTILAGE (HC) -39 PROTEIN
US-09-171-705-51

  Query Match             93.8%;  Score 61;  DB 3;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 0.00011;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VGYDDQESVKSK 12
              ||||||||||||
Db          5 VGYDDQESVKSK 16


RESULT 4
US-09-459-749D-17
; Sequence 17, Application US/09459749D
; Patent No. 6464975
; GENERAL INFORMATION:
;  APPLICANT: Millis, Albert J. T.
;  TITLE OF INVENTION: Compositions and Methods For Altering Cell Migration
;  FILE REFERENCE: 0794.016A
;  CURRENT APPLICATION NUMBER: US/09/459,749D
;  CURRENT FILING DATE:  1999-12-10
;  PRIOR APPLICATION NUMBER: 60/111,856
;  PRIOR FILING DATE: 1998-12-11
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
;   LENGTH: 383
;   TYPE: PRT
;   ORGANISM: Sus scrofa
US-09-459-749D-17

  Query Match             89.2%;  Score 58;  DB 4;  Length 383;
  Best Local Similarity   91.7%;  Pred. No. 0.011;
  Matches   11;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VGYDDQESVKSK 12
              |||:||||||:|
Db        326 VGYDDQESVKNK 337


RESULT 5
US-09-171-705-52
; Sequence 52, Application US/09171705
; Patent No. 6184204
; GENERAL INFORMATION:
;  APPLICANT: BOOTS, ANNA M.H.
;  APPLICANT: VERHEIJDEN, GILBERTUS F.M.
;  TITLE OF INVENTION: NOVEL PEPTIDES SUITABLE FOR USE IN ANTIGEN SPECIFIC
;  TITLE OF INVENTION: IMMUNOSUPPRESSIVE THERAPY
;  FILE REFERENCE: O/96198 US
;  CURRENT APPLICATION NUMBER: US/09/171,705
;  CURRENT FILING DATE: 1999-02-09
;  NUMBER OF SEQ ID NOS: 78
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 52
;   LENGTH: 16
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: DERIVED FROM
;   OTHER INFORMATION: SEQUENCE OF HUMAN CARTILAGE (HC) -39 PROTEIN
US-09-171-705-52

  Query Match             84.6%;  Score 55;  DB 3;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 0.0013;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          3 YDDQESVKSKV 13
              |||||||||||
Db          1 YDDQESVKSKV 11


RESULT 6
US-09-503-922-1
; Sequence 1, Application US/09503922
; Patent No. 6410706
; GENERAL INFORMATION:
;  APPLICANT: PAI, Hyun-Sook
;  APPLICANT: LIU, Jang-Ryol
;  APPLICANT: CHO, Hye-Sun
;  APPLICANT: KIM, Youn-Sung
;  TITLE OF INVENTION: A NOVEL CHITIN-BINDING RECEPTOR KINASE AND THE GENE
;  TITLE OF INVENTION: THEREOF
;  FILE REFERENCE: seq
;  CURRENT APPLICATION NUMBER: US/09/503,922
;  CURRENT FILING DATE: 2000-02-14
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 739
;   TYPE: PRT
;   ORGANISM: Nicotiana tabacum
US-09-503-922-1

  Query Match             72.3%;  Score 47;  DB 4;  Length 739;
  Best Local Similarity   53.8%;  Pred. No. 2.1;
  Matches    7;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 VGYDDQESVKSKV 13
              :|:|| |:::|:|
Db        306 IGYDDVEAIRIKV 318


RESULT 7
US-08-619-645-3
; Sequence 3, Application US/08619645
; Patent No. 5736507
;  GENERAL INFORMATION:
;    APPLICANT:  Boots, Anna M.H.
;    APPLICANT:  Verheijden, Gijsbertus F.M.
;    TITLE OF INVENTION:  No. 5736507el peptides derived from
;    TITLE OF INVENTION:  autoantigen for use in immunotherapy of autoimmune
;    TITLE OF INVENTION:  disease
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Akzo No. 5736507el Patent Department
;      STREET:  1300 Piccard Drive, Suite 206
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  U.S.A.
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25(EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/619,645
;      FILING DATE:  25-MAR-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Gormley, Mary E.
;      REGISTRATION NUMBER:  34,409
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 947-4433
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-619-645-3

  Query Match             70.8%;  Score 46;  DB 1;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 YDDQESVKS 11
              |||||||||
Db          1 YDDQESVKS 9


RESULT 8
US-08-634-493-3
; Sequence 3, Application US/08634493
; Patent No. 5843449
;  GENERAL INFORMATION:
;    APPLICANT:  A.M.H. Boots
;    APPLICANT:  G.F.M. Verheijden
;    APPLICANT:  E.S. Bos
;    TITLE OF INVENTION:  No. 5843449el Peptides derived from autoantigen for use
;    TITLE OF INVENTION:  in Immunotherapy of Autoimmune Diseases
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Akzo No. 5843449el Patent Department
;      STREET:  1300 Piccard Drive, Suite 206
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/634,493
;      FILING DATE:  18-APR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/619,645
;      FILING DATE:  25-MAR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/EP95/04201
;      FILING DATE:  25-OCT-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  NL 942031287
;      FILING DATE:  27-OCT-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  NL 952008860
;      FILING DATE:  07-APR-1995
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 258-5200
;      TELEFAX:  (301) 977-0847
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-634-493-3

  Query Match             70.8%;  Score 46;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 YDDQESVKS 11
              |||||||||
Db          1 YDDQESVKS 9
```

```
RESULT 9
US-09-039-198A-14
; Sequence 14, Application US/09039198A
; Patent No. 6200951
;  GENERAL INFORMATION:
;    APPLICANT:  Gray, Patrick W.
;    APPLICANT:  Tjoelker, Larry W.
;    TITLE OF INVENTION:  CHITINASE CHITIN-BINDING FRAGMENTS
;    NUMBER OF SEQUENCES:  34
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  233 South Wacker Drive/6300 Sears Tower
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/039,198A
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rin-Laures, Li-Hsien
;      REGISTRATION NUMBER:  33,547
;      REFERENCE/DOCKET NUMBER:  27866/34391
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 474-6300
;      TELEFAX:  (312) 474-0448
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  373 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-039-198A-14

  Query Match             70.8%;  Score 46;  DB 3;  Length 373;
  Best Local Similarity   69.2%;  Pred. No. 1.5;
  Matches    9;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 VGYDDQESVKSKV 13
              ||:|| || |:||
Db        311 VGFDDVESFKTKV 323


RESULT 10
US-09-039-198A-15
; Sequence 15, Application US/09039198A
; Patent No. 6200951
;  GENERAL INFORMATION:
;    APPLICANT:  Gray, Patrick W.
;    APPLICANT:  Tjoelker, Larry W.
;    TITLE OF INVENTION:  CHITINASE CHITIN-BINDING FRAGMENTS
;    NUMBER OF SEQUENCES:  34
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  233 South Wacker Drive/6300 Sears Tower
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/039,198A
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rin-Laures, Li-Hsien
;      REGISTRATION NUMBER:  33,547
;      REFERENCE/DOCKET NUMBER:  27866/34391
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 474-6300
;      TELEFAX:  (312) 474-0448
;  INFORMATION FOR SEQ ID NO:  15:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  373 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-039-198A-15

  Query Match             70.8%;  Score 46;  DB 3;  Length 373;
  Best Local Similarity   69.2%;  Pred. No. 1.5;
  Matches    9;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 VGYDDQESVKSKV 13
              ||:|| || |:||
Db        311 VGFDDVESFKTKV 323


RESULT 11
US-08-877-599-14
; Sequence 14, Application US/08877599
; Patent No. 6372212
;  GENERAL INFORMATION:
;    APPLICANT:  Gray, Patrick W.
;    TITLE OF INVENTION:  Chitinase Materials and Methods
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/877,599
;      FILING DATE:
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/663,618
;      FILING DATE:  14-JUN-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rin-Laures, Li-Hsien
;      REGISTRATION NUMBER:  33,547
;      REFERENCE/DOCKET NUMBER:  27866/33994
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  373 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-877-599-14

  Query Match             70.8%;  Score 46;  DB 4;  Length 373;
  Best Local Similarity   69.2%;  Pred. No. 1.5;
```

```
  Matches    9;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 VGYDDQESVKSKV 13
              |:|| || |:|:|
Db        311 VGFDDVESFKTKV 323


RESULT 12
US-08-877-599-15
; Sequence 15, Application US/08877599
; Patent No. 6372212
;  GENERAL INFORMATION:
;    APPLICANT:  Gray, Patrick W.
;    TITLE OF INVENTION:  Chitinase Materials and Methods
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/877,599
;      FILING DATE:
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/663,618
;      FILING DATE:  14-JUN-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rin-Laures, Li-Hsien
;      REGISTRATION NUMBER:  33,547
;      REFERENCE/DOCKET NUMBER:  27866/33994
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  15:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  373 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-877-599-15

  Query Match             70.8%;  Score 46;  DB 4;  Length 373;
  Best Local Similarity   69.2%;  Pred. No. 1.5;
  Matches    9;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 VGYDDQESVKSKV 13
              |:|| || |:|:|
Db        311 VGFDDVESFKTKV 323


RESULT 13
US-09-267-574-14
; Sequence 14, Application US/09267574
; Patent No. 6399571
; GENERAL INFORMATION:
;  APPLICANT: Gray, Patrick W.
;  APPLICANT: Tjoelker, Larry W.
;  TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
;  FILE REFERENCE: 27866/35407
;  CURRENT APPLICATION NUMBER: US/09/267,574
;  CURRENT FILING DATE: 1999-03-12
;  EARLIER APPLICATION NUMBER: 09/039,198
;  EARLIER FILING DATE: 1998-03-12
;  NUMBER OF SEQ ID NOS: 39
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
;   LENGTH: 373
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-267-574-14

  Query Match             70.8%;  Score 46;  DB 4;  Length 373;
  Best Local Similarity   69.2%;  Pred. No. 1.5;
  Matches    9;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 VGYDDQESVKSKV 13
              |:|| || |:|:|
Db        311 VGFDDVESFKTKV 323


RESULT 14
US-09-267-574-15
; Sequence 15, Application US/09267574
; Patent No. 6399571
; GENERAL INFORMATION:
;  APPLICANT: Gray, Patrick W.
;  APPLICANT: Tjoelker, Larry W.
;  TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
;  FILE REFERENCE: 27866/35407
;  CURRENT APPLICATION NUMBER: US/09/267,574
;  CURRENT FILING DATE: 1999-03-12
;  EARLIER APPLICATION NUMBER: 09/039,198
;  EARLIER FILING DATE: 1998-03-12
;  NUMBER OF SEQ ID NOS: 39
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
;   LENGTH: 373
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-267-574-15

  Query Match             70.8%;  Score 46;  DB 4;  Length 373;
  Best Local Similarity   69.2%;  Pred. No. 1.5;
  Matches    9;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 VGYDDQESVKSKV 13
              |:|| || |:|:|
Db        311 VGFDDVESFKTKV 323


RESULT 15
US-08-486-839-6
; Sequence 6, Application US/08486839
; Patent No. 5928928
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  A human chitinase, its recombinant
;    TITLE OF INVENTION:  production, its use for decomposing chitin, its use
;    TITLE OF INVENTION:  in therapy or prophylaxis against infection diseases.
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hoffmann & Baron
;      STREET:  350 Jericho Turnpike
;      CITY:  Jericho
;      STATE:  New York
;      COUNTRY:  United States of America
;      ZIP:  11756
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;      SOFTWARE:
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/486,839
;      FILING DATE:  07 - June - 1995
```

```
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Baron, Ronald J.
;       REGISTRATION NUMBER:  29,281
;       REFERENCE/DOCKET NUMBER:  294-28
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (516) 822-3550
;       TELEFAX:  (516) 822-3582
;   INFORMATION FOR SEQ ID NO:  6:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  387 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  unknown
;       TOPOLOGY:  unknown
;     MOLECULE TYPE:  protein
;     HYPOTHETICAL:  NO
US-08-486-839-6

  Query Match             70.8%;  Score 46;  DB 2;  Length 387;
  Best Local Similarity   69.2%;  Pred. No. 1.6;
  Matches    9;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 VGYDDQESVKSKV 13
              :|:|| |:| |:|:|
Db        332 VGFDDVESPKIKV 344


Search completed: September 24, 2003, 13:51:16
Job time : 16.7753 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 24, 2003, 13:47:40 ; Search time 29.0674 Seconds
(without alignments)
67.670 Million cell updates/sec

Title: US-09-744-282-7
Perfect score: 65
Sequence: 1 VGYDDQESVKSKV 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 566894 seqs, 151307093 residues

Total number of hits satisfying chosen parameters: 566894

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	65	100.0	383	15	US-10-097-340-45	Sequence 45, Appl
2	58	89.2	383	10	US-09-459-749D-17	Sequence 17, Appl
3	52	80.0	14	10	US-09-826-296-324	Sequence 324, App
4	52	80.0	14	11	US-09-791-393-109	Sequence 109, App
5	52	80.0	14	12	US-09-791-389-109	Sequence 109, App
6	46	70.8	373	12	US-10-161-547-14	Sequence 14, Appl
7	46	70.8	373	12	US-10-161-547-15	Sequence 15, Appl
8	46	70.8	445	15	US-10-004-219B-10	Sequence 10, Appl
9	46	70.8	466	12	US-10-161-547-2	Sequence 2, Appli
10	46	70.8	466	12	US-10-161-547-4	Sequence 4, Appli
11	43	66.2	371	10	US-09-923-844B-2	Sequence 2, Appli
12	41	63.1	384	15	US-10-194-489-21	Sequence 21, Appl
13	41	63.1	392	15	US-10-194-489-22	Sequence 22, Appl
14	41	63.1	400	15	US-10-194-489-23	Sequence 23, Appl
15	41	63.1	411	15	US-10-194-489-1	Sequence 1, Appli
16	41	63.1	411	15	US-10-194-489-3	Sequence 3, Appli
17	41	63.1	411	15	US-10-194-489-4	Sequence 4, Appli
18	41	63.1	411	15	US-10-194-489-5	Sequence 5, Appli
19	41	63.1	411	15	US-10-194-489-6	Sequence 6, Appli
20	41	63.1	411	15	US-10-194-489-7	Sequence 7, Appli
21	41	63.1	411	15	US-10-194-489-8	Sequence 8, Appli
22	41	63.1	411	15	US-10-194-489-9	Sequence 9, Appli
23	41	63.1	411	15	US-10-194-489-10	Sequence 10, Appl
24	41	63.1	411	15	US-10-194-489-11	Sequence 11, Appl
25	41	63.1	411	15	US-10-194-489-12	Sequence 12, Appl
26	41	63.1	411	15	US-10-194-489-13	Sequence 13, Appl
27	41	63.1	411	15	US-10-194-489-14	Sequence 14, Appl
28	41	63.1	411	15	US-10-194-489-15	Sequence 15, Appl
29	41	63.1	411	15	US-10-194-489-16	Sequence 16, Appl
30	41	63.1	411	15	US-10-194-489-17	Sequence 17, Appl
31	41	63.1	411	15	US-10-194-489-18	Sequence 18, Appl
32	41	63.1	411	15	US-10-194-489-19	Sequence 19, Appl
33	41	63.1	411	15	US-10-194-489-20	Sequence 20, Appl
34	38	58.5	490	15	US-10-218-743-41	Sequence 41, Appl
35	38	58.5	509	15	US-10-218-743-35	Sequence 35, Appl
36	38	58.5	509	15	US-10-218-743-38	Sequence 38, Appl
37	38	58.5	536	15	US-10-218-743-21	Sequence 21, Appl
38	38	58.5	555	15	US-10-218-743-15	Sequence 15, Appl
39	38	58.5	555	15	US-10-218-743-18	Sequence 18, Appl
40	37	56.9	309	15	US-10-068-374-5	Sequence 5, Appli
41	37	56.9	400	11	US-09-903-185-2	Sequence 2, Appli
42	37	56.9	462	11	US-09-903-185-13	Sequence 13, Appl
43	37	56.9	467	10	US-09-863-547B-1	Sequence 1, Appli
44	37	56.9	467	15	US-10-068-374-2	Sequence 2, Appli
45	36	55.4	852	9	US-09-815-242-11535	Sequence 11535, A

ALIGNMENTS

RESULT 1
US-10-097-340-45
; Sequence 45, Application US/10097340
; Publication No. US20030087250A1
; GENERAL INFORMATION:
; APPLICANT: John MONAHAN
; APPLICANT: Manjula GANNAVARAPU
; APPLICANT: Sebastian HOERSCH
; APPLICANT: Shubhangi KAMATKAR
; APPLICANT: Steve G. KOVATS
; APPLICANT: Rachel E. MEYERS
; APPLICANT: Michael MORRISEY
; APPLICANT: Peter OLANDT
; APPLICANT: Ami SEN
; APPLICANT: Peter VEIBY
; APPLICANT: Gordon B. MILLS
; APPLICANT: Robert C. BAST, Jr.
; APPLICANT: Karen LU
; APPLICANT: Rosemarie SCHMANDT
; APPLICANT: Xumei ZHAO
; APPLICANT: Karen GLATT
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
; TITLE OF INVENTION: Assessment, Prevention, and Therapy of Ovarian Cancer
; FILE REFERENCE: MRI-030
; CURRENT APPLICATION NUMBER: US/10/097,340
; CURRENT FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 60/275,025
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/325,149
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/276,026
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/324,967
; PRIOR FILING DATE: 2001/09/26
; PRIOR APPLICATION NUMBER: 60/311,732
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/325,102
; PRIOR FILING DATE: 2001-09-26

; PRIOR APPLICATION NUMBER: 60/323,580
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: PastSEQ for Windows Version 4.0
; SEQ ID NO 45
; LENGTH: 383
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-097-340-45

Query Match 100.0%; Score 65; DB 15; Length 383;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VGYDDQESVKSKV 13
||||||||||||| |
Db 326 VGYDDQESVKSKV 338

RESULT 2
US-09-459-749D-17
; Sequence 17, Application US/09459749D
; Patent No. US20020136716A1
; GENERAL INFORMATION:
; APPLICANT: Millis, Albert J. T.
; TITLE OF INVENTION: Compositions and Methods For Altering Cell Migration
; FILE REFERENCE: 0794.016A
; CURRENT APPLICATION NUMBER: US/09/459,749D
; CURRENT FILING DATE: 1999-12-10
; PRIOR APPLICATION NUMBER: 60/111,856
; PRIOR FILING DATE: 1998-12-11
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 383
; TYPE: PRT
; ORGANISM: Sus scrofa
US-09-459-749D-17

Query Match 89.2%; Score 58; DB 10; Length 383;
Best Local Similarity 91.7%; Pred. No. 0.03;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VGYDDQESVKSK 12
|||||||||| :|
Db 326 VGYDDQESVKNK 337

RESULT 3
US-09-826-290-324
; Sequence 324, Application US/09826290
; Patent No. US20020164668A1
; GENERAL INFORMATION:
; APPLICANT: Durham, L.Kathryn
; APPLICANT: Friedman, David L.
; APPLICANT: Herath, Herath Mudiyanselage Athula Chandrasiri
; APPLICANT: Kimmel, Lida H.
; APPLICANT: Parekh, Rajesh Bhikhu
; APPLICANT: Potter, David M.
; APPLICANT: Rohlff, Christian
; APPLICANT: Silber, B. Michael
; APPLICANT: Stiger, Thomas R.
; APPLICANT: Sunderland, P. Trey
; APPLICANT: Townsend, Robert Reid
; APPLICANT: White, Frost
; APPLICANT: Williams, Stephen A.
; TITLE OF INVENTION: Nucleic Acid Molecules, Polypeptides and
; TITLE OF INVENTION: Uses Therefor, Including Diagnosis and Treatment of
; TITLE OF INVENTION: Alzheimer's Disease
; FILE REFERENCE: 2572-1-001 N2
; CURRENT APPLICATION NUMBER: US/09/826,290
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: US 60/194,504
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: US 60/253,647
; PRIOR FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 492
; SOFTWARE: PastSEQ for Windows Version 4.0
; SEQ ID NO 324
; LENGTH: 14
; TYPE: PRT
; ORGANISM: homo sapien
US-09-826-290-324

Query Match 80.0%; Score 52; DB 10; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.61;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VGYDDQESVK 10
||||||||||
Db 5 VGYDDQESVK 14

RESULT 4
US-09-791-393-109
; Sequence 109, Application US/09791393
; Publication No. US20030032200A1
; GENERAL INFORMATION:
; APPLICANT: Herath, Herath Mudiyanselage Athula Chandrasiri
; APPLICANT: Parekh, Rajesh Bhikhu
; APPLICANT: Rohlff, Christian
; TITLE OF INVENTION: Proteins, Genes and Their Use for
; TITLE OF INVENTION: Diagnosis and Treatment of Bipolar Affective Disorder (BAD)
; TITLE OF INVENTION: and Unipolar Depression
; FILE REFERENCE: 2543-1-001 N1
; CURRENT APPLICATION NUMBER: US/09/791,393
; CURRENT FILING DATE: 2002-01-02
; EARLIER APPLICATION NUMBER: GB 0004412.3
; EARLIER FILING DATE: 2000-02-24
; EARLIER APPLICATION NUMBER: GB 0030050.9
; EARLIER FILING DATE: 2000-12-08
; EARLIER APPLICATION NUMBER: US 60/254,830
; EARLIER FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 308
; SOFTWARE: PastSEQ for Windows Version 4.0
; SEQ ID NO 109
; LENGTH: 14
; TYPE: PRT
; ORGANISM: homo sapien
US-09-791-393-109

Query Match 80.0%; Score 52; DB 11; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.61;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VGYDDQESVK 10
||||||||||
Db 5 VGYDDQESVK 14

RESULT 5
US-09-791-389-109
; Sequence 109, Application US/09791389
; Publication No. US20030032773A1
; GENERAL INFORMATION:
; APPLICANT: Herath, Herath Mudiyanselage Athula Chandrasiri
; APPLICANT: Parekh, Rajesh Bhikhu
; APPLICANT: Rohlff, Christian
; APPLICANT: Terrett, Jonathan Alexander
; APPLICANT: Tyson, Kerry Louise
; TITLE OF INVENTION: Proteins, Genes and Their Use for
; TITLE OF INVENTION: Diagnosis and Treatment of Bipolar Affective Disorder (BAD)
; TITLE OF INVENTION: and Unipolar Depression
; FILE REFERENCE: 2543-1-001 N2
; CURRENT APPLICATION NUMBER: US/09/791,389
; CURRENT FILING DATE: 2001-02-23

```
;  PRIOR APPLICATION NUMBER: GB 0004412.3
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: GB 0030050.9
;  PRIOR FILING DATE: 2000-12-08
;  PRIOR APPLICATION NUMBER: US 60/254,830
;  PRIOR FILING DATE: 2000-12-12
;  NUMBER OF SEQ ID NOS: 308
;  SOFTWARE: PastSEQ for Windows Version 4.0
; SEQ ID NO 109
;   LENGTH: 14
;   TYPE: PRT
;   ORGANISM: homo sapien
US-09-791-389-109

  Query Match           80.0%;  Score 52;  DB 11;  Length 14;
  Best Local Similarity  100.0%;  Pred. No. 0.01;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    2;  Gaps    2;

Qy          1 VGYDDQESVK 10
              ||||||||||
Db          5 VGYDDQESVK 14


RESULT 6
US-10-161-547-14
; Sequence 14, Application US/10161547
; Publication No. US20030143216A1
; GENERAL INFORMATION:
;  APPLICANT: Gray, Patrick W.
;  APPLICANT:  Tjoelker, Larry W.
;  TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
;  FILE REFERENCE: 27866/35407
;  CURRENT APPLICATION NUMBER: US/10/161,547
;  CURRENT FILING DATE:  2002-06-03
;  PRIOR APPLICATION NUMBER: US/09/267,574
;  PRIOR FILING DATE: 1999-03-12
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/039,198
;  PRIOR FILING DATE: EARLIER FILING DATE: 1998-03-12
;  NUMBER OF SEQ ID NOS: 39
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
;   LENGTH: 373
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-161-547-14

  Query Match           70.8%;  Score 46;  DB 12;  Length 373;
  Best Local Similarity  69.2%;  Pred. No. 3.7;
  Matches    9;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 VGYDDQESVKSKV 13
              ||:||:|| |:||
Db        311 VGFDDVESFKTKV 323


RESULT 7
US-10-161-547-15
; Sequence 15, Application US/10161547
; Publication No. US20030143216A1
; GENERAL INFORMATION:
;  APPLICANT: Gray, Patrick W.
;  APPLICANT:  Tjoelker, Larry W.
;  TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
;  FILE REFERENCE: 27866/35407
;  CURRENT APPLICATION NUMBER: US/10/161,547
;  CURRENT FILING DATE:  2002-06-03
;  PRIOR APPLICATION NUMBER: US/09/267,574
;  PRIOR FILING DATE: 1999-03-12
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/039,198
;  PRIOR FILING DATE: EARLIER FILING DATE: 1998-03-12
;  NUMBER OF SEQ ID NOS: 39
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
;   LENGTH: 373
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-161-547-15

  Query Match           70.8%;  Score 46;  DB 12;  Length 373;
  Best Local Similarity  69.2%;  Pred. No. 3.7;
  Matches    9;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 VGYDDQESVKSKV 13
              ||:||:|| |:||
Db        311 VGFDDVESFKIKV 323


RESULT 8
US-10-004-219B-10
; Sequence 10, Application US/10004219B
; Publication No. US20030087414A1
; GENERAL INFORMATION:
;  APPLICANT: Macrozyme
;  APPLICANT:  Aerts, Johannes M.F.G.
;  APPLICANT:  Boot, Rolf G.
;  TITLE OF INVENTION: A mammalian mucinase, its recombinant production, and
;  TITLE OF INVENTION:  its use in therapy or prophylaxis against diseases in
;  TITLE OF INVENTION:  which mucus is involved or infection diseases
;  FILE REFERENCE: 2183-5136US
;  CURRENT APPLICATION NUMBER: US/10/004,219B
;  CURRENT FILING DATE:  2001-11-02
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
;   LENGTH: 445
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (1)..(445)
;   OTHER INFORMATION: /note="Human chitotriosidase"
US-10-004-219B-10

  Query Match           70.8%;  Score 46;  DB 15;  Length 445;
  Best Local Similarity  69.2%;  Pred. No. 4.4;
  Matches    9;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 VGYDDQESVKSKV 13
              ||:|| |:||
Db        311 VGFDDVESFKTKV 323


RESULT 9
US-10-161-547-2
; Sequence 2, Application US/10161547
; Publication No. US20030143216A1
; GENERAL INFORMATION:
;  APPLICANT: Gray, Patrick W.
;  APPLICANT:  Tjoelker, Larry W.
;  TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
;  FILE REFERENCE: 27866/35407
;  CURRENT APPLICATION NUMBER: US/10/161,547
;  CURRENT FILING DATE:  2002-06-03
;  PRIOR APPLICATION NUMBER: US/09/267,574
;  PRIOR FILING DATE: 1999-03-12
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/039,198
;  PRIOR FILING DATE: EARLIER FILING DATE: 1998-03-12
;  NUMBER OF SEQ ID NOS: 39
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 466
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-161-547-2

  Query Match           70.8%;  Score 46;  DB 12;  Length 466;
```

```
  Best Local Similarity   69.2%;  Pred. No. 4.7;
  Matches    9;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 VGYDDQESVKSKV 13
              ||:|| || |:.|
Db        332 VGFDDVESFKTKV 344


RESULT 10
US-10-161-547-4
; Sequence 4, Application US/10161547
; Publication No. US20030143216A1
; GENERAL INFORMATION:
;  APPLICANT: Gray, Patrick W.
;  APPLICANT:  Tjoelker, Larry W.
;  TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
;  FILE REFERENCE: 27866/35407
;  CURRENT APPLICATION NUMBER: US/10/161,547
;  CURRENT FILING DATE:  2002-06-03
;  PRIOR APPLICATION NUMBER: US/09/267,574
;  PRIOR FILING DATE: 1999-03-12
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/039,198
;  PRIOR FILING DATE: EARLIER FILING DATE: 1998-03-12
;  NUMBER OF SEQ ID NOS: 39
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
;   LENGTH: 466
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-161-547-4

  Query Match             70.8%;  Score 46;  DB 12;  Length 466;
  Best Local Similarity   69.2%;  Pred. No. 4.7;
  Matches    9;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 VGYDDQESVKSKV 13
              :|:|: || |:|:
Db        332 VGFDDVESFKTKV 344


RESULT 11
US-09-923-844B-2
; Sequence 2, Application US/09923844B
; Patent No. US20020166143A1
; GENERAL INFORMATION:
;  APPLICANT: Pioneer Hi-Bred International, Inc.
;  APPLICANT:  Bao, Zhongmeng
;  APPLICANT:  Lu, Guihua
;  TITLE OF INVENTION: Sclerotinia-Inducible Genes and
;  TITLE OF INVENTION:  Promoters and Their Uses
;  FILE REFERENCE: 35718/234631
;  CURRENT APPLICATION NUMBER: US/09/923,844B
;  CURRENT FILING DATE:  2001-08-07
;  PRIOR APPLICATION NUMBER: US 60/224,603
;  PRIOR FILING DATE: 2000-08-11
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 371
;   TYPE: PRT
;   ORGANISM: Helianthus annuus
US-09-923-844B-2

  Query Match             66.2%;  Score 43;  DB 10;  Length 371;
  Best Local Similarity   61.5%;  Pred. No. 12;
  Matches    8;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 VGYDDQESVKSKV 13
              :|||:  ||: :|
Db        325 IGYDDVNSVQRKV 337


RESULT 12
```

```
US-10-194-489-21
; Sequence 21, Application US/10194489
; Publication No. US20030086946A1
; GENERAL INFORMATION:
;  APPLICANT: Tweten, Rodney K.
;  APPLICANT:  Melton, Jody
;  TITLE OF INVENTION: Antigenic, No. US20030086946A1-Toxic Mutants of Clostridium s
;  TITLE OF INVENTION:  Vaccines, Antibodies, Sera, and Methods of Treatment Therewi
;  FILE REFERENCE: 5864.020
;  CURRENT APPLICATION NUMBER: US/10/194,489
;  CURRENT FILING DATE:  2002-07-10
;  PRIOR APPLICATION NUMBER: 60/305,725
;  PRIOR FILING DATE: 2001-07-16
;  NUMBER OF SEQ ID NOS: 23
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 21
;   LENGTH: 384
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: Completely synthesized.
US-10-194-489-21

  Query Match             63.1%;  Score 41;  DB 15;  Length 384;
  Best Local Similarity   50.0%;  Pred. No. 28;
  Matches    6;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 GYDDQESVKSKV 13
              |:|| | :|:|:
Db         21 GYEDNEDLKAKI 32


RESULT 13
US-10-194-489-22
; Sequence 22, Application US/10194489
; Publication No. US20030086946A1
; GENERAL INFORMATION:
;  APPLICANT: Tweten, Rodney K.
;  APPLICANT:  Melton, Jody
;  TITLE OF INVENTION: Antigenic, No. US20030086946A1-Toxic Mutants of Clostridium s
;  TITLE OF INVENTION:  Vaccines, Antibodies, Sera, and Methods of Treatment Therewi
;  FILE REFERENCE: 5864.020
;  CURRENT APPLICATION NUMBER: US/10/194,489
;  CURRENT FILING DATE:  2002-07-10
;  PRIOR APPLICATION NUMBER: 60/305,725
;  PRIOR FILING DATE: 2001-07-16
;  NUMBER OF SEQ ID NOS: 23
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 22
;   LENGTH: 392
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: Completely synthesized.
US-10-194-489-22

  Query Match             63.1%;  Score 41;  DB 15;  Length 392;
  Best Local Similarity   50.0%;  Pred. No. 29;
  Matches    6;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 GYDDQESVKSKV 13
              |:|| | :|:|:
Db         21 GYEDNEDLKAKI 32


RESULT 14
US-10-194-489-23
; Sequence 23, Application US/10194489
; Publication No. US20030086946A1
; GENERAL INFORMATION:
;  APPLICANT: Tweten, Rodney K.
;  APPLICANT:  Melton, Jody
;  TITLE OF INVENTION: Antigenic, No. US20030086946A1-Toxic Mutants of Clostridium s
```

```
;  TITLE OF INVENTION:  Vaccines, Antibodies, Sera, and Methods of Treatment Therewith
;  FILE REFERENCE: 5864.020
;  CURRENT APPLICATION NUMBER: US/10/194,489
;  CURRENT FILING DATE:  2002-07-10
;  PRIOR APPLICATION NUMBER: 60/305,725
;  PRIOR FILING DATE: 2001-07-16
;  NUMBER OF SEQ ID NOS: 23
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 23
;   LENGTH: 400
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: Completely synthesized.
US-10-194-489-23

  Query Match             63.1%;  Score 41;  DB 15;  Length 400;
  Best Local Similarity   50.0%;  Pred. No. 30;
  Matches    6;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 GYDDQESVKSKV 13
              ||:  .. :|:.:
Db         21 GYEDNEDLKAKI 32


RESULT 15
US-10-194-489-1
; Sequence 1, Application US/10194489
; Publication No. US20030086946A1
; GENERAL INFORMATION:
;  APPLICANT: Tweten, Rodney K.
;  APPLICANT:  Melton, Jody
;  TITLE OF INVENTION: Antigenic, No. US20030086946A1-Toxin Mutants of Clostridium sept
;  TITLE OF INVENTION:  Vaccines, Antibodies, Sera, and Methods of Treatment Therewith
;  FILE REFERENCE: 5864.020
;  CURRENT APPLICATION NUMBER: US/10/194,489
;  CURRENT FILING DATE:  2002-07-10
;  PRIOR APPLICATION NUMBER: 60/305,725
;  PRIOR FILING DATE: 2001-07-16
;  NUMBER OF SEQ ID NOS: 23
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 411
;   TYPE: PRT
;   ORGANISM: Clostridium septicum
US-10-194-489-1

  Query Match             63.1%;  Score 41;  DB 15;  Length 411;
  Best Local Similarity   50.0%;  Pred. No. 30;
  Matches    6;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 GYDDQESVKSKV 13
              ||:| | ::.:
Db         21 GYEDNEDLKAKI 32


Search completed: September 24, 2003, 14:22:30
Job time : 30.0674 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 24, 2003, 13:34:08 ; Search time 15.7753 Seconds
(without alignments)
79.250 Million cell updates/sec

Title: US-09-744-282-7
Perfect score: 65
Sequence: 1 VGYDDQESVKSKV 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 98168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	65	100.0	383	2	A49562	cartilage glycopro
2	58	89.2	383	2	S51327	heparin-binding gl
3	54	83.1	405	2	S61551	breast-regressing
4	48	73.8	371	2	T04755	chitinase homolog
5	46	70.8	421	2	T04753	hypothetical prote
6	45	69.2	379	2	T04762	chitinase homolog
7	44	67.7	248	2	T04758	hypothetical prote
8	44	67.7	332	2	T04754	hypothetical prote
9	43	66.2	366	2	T04763	chitinase homolog
10	43	66.2	378	2	S51591	chitinase (EC 3.2.
11	43	66.2	525	2	T44445	chitinase (EC 3.2.
12	42	64.6	365	2	T04757	chitinase homolog
13	42	64.6	526	2	H81312	protein-export mem
14	42	64.6	633	2	T24898	hypothetical prote
15	40	61.5	349	2	G90165	conserved hypothet
16	40	61.5	358	2	T00578	probable GDSL-moti
17	40	61.5	359	2	H84706	probable GDSL-moti
18	40	61.5	398	2	T04761	chitinase homolog
19	40	61.5	588	2	S42074	aspartic proteinas
20	40	61.5	3006	2	T28625	variant-specific s
21	39	60.0	1545	2	T14285	DNA (cytosine-5-)-
22	38	58.5	373	2	S15161	long-chain-fatty-a
23	38	58.5	486	2	I40548	bifunctional cellu
24	38	58.5	499	2	A27198	cellulase (EC 3.2.
25	38	58.5	508	2	G59593	cellulase (EC 3.2.
26	38	58.5	508	2	A26874	cellulase (EC 3.2.
27	38	58.5	609	2	T28896	hypothetical prote
28	38	58.5	1022	2	T51257	calmodulin-binding
29	38	58.5	1022	2	T50928	calmodulin-binding
30	37	56.9	364	1	JE0292	fungal stress prot
31	37	56.9	504	2	A38221	chitinase (EC 3.2.
32	37	56.9	554	2	A56596	chitinase (EC 3.2.
33	37	56.9	718	2	T29706	hypothetical prote
34	37	56.9	4063	2	T42993	probable spectrin
35	37	56.9	4101	2	T23630	hypothetical prote
36	36.5	56.2	1005	2	F81305	probable integral
37	36	55.4	176	2	E72348	conserved hypothet
38	36	55.4	253	2	E69098	phosphate transpor
39	36	55.4	350	2	B95957	probable transcrip
40	36	55.4	370	2	E42951	long-chain-fatty-a
41	36	55.4	486	2	AB1938	glutamyl-tRNA (Gln
42	36	55.4	499	2	S52422	chitinase (EC 3.2.
43	36	55.4	499	2	S04856	chitinase (EC 3.2.
44	36	55.4	511	2	S61166	probable membrane
45	36	55.4	572	2	G84238	hypothetical prote

ALIGNMENTS

RESULT 1
A49562
cartilage glycoprotein gp39 precursor - human
N;Alternate names: 39K synovial protein
C;Species: Homo sapiens (man)
C;Date: 23-Mar-1995 #sequence_revision 23-Mar-1995 #text_change 22-Jun-1999
C;Accession: A49562; S10677; A35262
R;Hakala, B.E.; White, C.; Recklies, A.D.
J. Biol. Chem. 268, 25803-25810, 199
A;Title: Human cartilage gp-39, a major secretory product of articular chondrocytes
A;Reference number: A49562; MUID:94064658; PMID:8245017
A;Accession: A49562
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-383 <HAK>
A;Cross-references: GB:M80927; NID:g348911; PIDN:AAA16074.1; PID:g348912
R;Nyirkos, P.; Golds, E.E.
Biochem. J. 269, 265-268, 1990
A;Title: Human synovial cells secrete a 39 kDa protein similar to a bovine mammary
A;Reference number: S10677; MUID:90328983; PMID:2375755
A;Accession: S10677
A;Molecule type: protein
A;Residues: 22-40,'X',42-45 <NY2>
C;Superfamily: Streptomyces chitinase chi4C
C;Keywords: cartilage; extracellular protein; glycoprotein
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-383/Product: cartilage glycoprotein gp39 #status predicted <MAT>

Query Match 100.0%; Score 65; DB 2; Length 383;
Best Local Similarity 100.0%; Pred. No. 0.00069;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 VGYDDQESVKSKV 13
              |||||||||||||
Db        326 VGYDDQESVKSKV 338
```

RESULT 2
S51327
heparin-binding glycoprotein 38K - pig
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 22-Jun-1999
C;Accession: S51327
R;Shackelton, L.M.; Mann, D.M.; Millis, A.J.T.
submitted to the EMBL Data Library, January 1995
A;Description: Identification of a 38kDa heparin-binding glycoprotein (gp38k) in di
A;Reference number: S51327
A;Accession: S51327
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-383 <SHA>
A;Cross-references: EMBL:Z47803; NID:g634097; PIDN:CAA87764.1; PID:g634098

```
C;Superfamily: Streptomyces chitinase chi40

  Query Match             89.2%;  Score 58;  DB 2;  Length 383;
  Best Local Similarity   91.7%;  Pred. No. 0.013;
  Matches   11;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VGYDDQESVKSK 12
              ||:|||||||:|
Db        326 VGYDDQESVKNK 337


RESULT 3
S61551
breast-regressing protein brp39 precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 26-Jun-2003
C;Accession: S61551; S61550; I48271
R;Morrison, B.W.; Leder, P.
Oncogene 9, 3417-3426, 1994
A;Title: neu and ras initiate murine mammary tumors that share genetic markers generally
A;Reference number: I48271; MUID:95060797; PMID:7970709
A;Accession: S61551
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-405 <MOR1>
A;Cross-references: EMBL:X93035; NID:g1085065; PIDN:CAA63603.1; PID:g1085066
R;Morrison, B.W.
submitted to the EMBL Data Library, November 1995
A;Reference number: S61550
A;Accession: S61550
A;Molecule type: mRNA
A;Residues: 1-245,'I',247-330,'H',332-350,'MVWALDLDDFQGTCQPKEFFPLTNAIKDALA' <MOR2>
A;Cross-references: EMBL:X93035; NID:g1085065; PIDN:CAA63603.1; PID:g1085066
A;Note: the differences at the carboxyl end are due to a frameshift error
C;Genetics:
A;Gene: brp39
C;Superfamily: Streptomyces chitinase chi40
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-405/Product: breast-regressing protein brp39 #status predicted <MAT>

  Query Match             83.1%;  Score 54;  DB 2;  Length 405;
  Best Local Similarity   76.9%;  Pred. No. 0.07;
  Matches   10;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VGYDDQESVKSKV 13
              |||:|:||||:||
Db        327 VGYEDKESVKNKV 339


RESULT 4
T04756
chitinase homolog T16H5.110 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-Jun-1999
C;Accession: T04756
R;Bevan, M.; De Haan, M.; Maarse, A.C.; Grivell, L.A.; Bancroft, I.; Mewes, H.W.; Mayer,
submitted to the Protein Sequence Database, June 1998
A;Reference number: Z15383
A;Accession: T04756
A;Molecule type: DNA
A;Residues: 1-371 <BEV>
A;Cross-references: EMBL:AL024486
A;Experimental source: cultivar Columbia; BAC clone T16H5
C;Genetics:
A;Map position: 4
A;Introns: 221/1; 351/1
A;Note: T16H5.110
C;Superfamily: Streptomyces chitinase chi40

  Query Match             73.8%;  Score 48;  DB 2;  Length 371;
  Best Local Similarity   61.5%;  Pred. No. 0.77;
  Matches    8;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 VGYDDQESVKSKV 13
              :|||| :|: |||
Db        311 IGYDDNQSIVSKV 323


RESULT 5
T04753
hypothetical protein T16H5.80 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 26-Aug-1999
C;Accession: T04753
R;Bevan, M.; De Haan, M.; Maarse, A.C.; Grivell, L.A.; Bancroft, I.; Mewes, H.W.; Ma
submitted to the Protein Sequence Database, June 1998
A;Reference number: Z15383
A;Accession: T04753
A;Molecule type: DNA
A;Residues: 1-421 <BEV>
A;Cross-references: EMBL:AL024486
A;Experimental source: cultivar Columbia; BAC clone T16H5
C;Genetics:
A;Map position: 4
A;Introns: 278/1
A;Note: T16H5.80
C;Superfamily: Streptomyces chitinase chi40

  Query Match             70.8%;  Score 46;  DB 2;  Length 421;
  Best Local Similarity   61.5%;  Pred. No. 2;
  Matches    8;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 VGYDDQESVKSKV 13
              :|||| :|| :|'
Db        364 IGYDDHQSVVAKV 376


RESULT 6
T04762
chitinase homolog T16H5.170 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-Jun-1999
C;Accession: T04762
R;Bevan, M.; De Haan, M.; Maarse, A.C.; Grivell, L.A.; Bancroft, I.; Mewes, H.W.; Ma
submitted to the Protein Sequence Database, June 1998
A;Reference number: Z15383
A;Accession: T04762
A;Molecule type: DNA
A;Residues: 1-379 <BEV>
A;Cross-references: EMBL:AL024486
A;Experimental source: cultivar Columbia; BAC clone T16H5
C;Genetics:
A;Map position: 4
A;Introns: 232/1; 362/1
A;Note: T16H5.170
C;Superfamily: Streptomyces chitinase chi40

  Query Match             69.2%;  Score 45;  DB 2;  Length 379;
  Best Local Similarity   53.8%;  Pred. No. 2.7;
  Matches    7;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 VGYDDQESVKSKV 13
              :|||| :|: :||
Db        322 IGYDDNQSIVTKV 334


RESULT 7
T04758
hypothetical protein T16H5.130 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-Jun-1999
C;Accession: T04758
R;Bevan, M.; De Haan, M.; Maarse, A.C.; Grivell, L.A.; Bancroft, I.; Mewes, H.W.; Ma
submitted to the Protein Sequence Database, June 1998
A;Reference number: Z15383
A;Accession: T04758
```

A;Molecule type: DNA
A;Residues: 1-248 <BEV>
A;Cross-references: EMBL:AL024486
A;Experimental source: cultivar Columbia; BAC clone T16H5
C;Genetics:
A;Map position: 4
A;Introns: 112/1; 242/1
A;Note: T16H5.130
C;Superfamily: Streptomyces chitinase chi40

  Query Match 67.7%; Score 44; DB 2; Length 248;
  Best Local Similarity 53.8%; Pred. No. 2.6;
  Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 VGYDDQESVKSKV 13
Db 202 IGYDSEESIVTKV 214

RESULT 8
T04754
hypothetical protein T16H5.90 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-Jun-1999
C;Accession: T04754
R;Bevan, M.; De Haan, M.; Maarse, A.C.; Grivell, L.A.; Bancroft, I.; Mewes, H.W.; Mayer,
submitted to the Protein Sequence Database, June 1998
A;Reference number: Z15383
A;Accession: T04754
A;Molecule type: DNA
A;Residues: 1-332 <BEV>
A;Cross-references: EMBL:AL024486
A;Experimental source: cultivar Columbia; BAC clone T16H5
C;Genetics:
A;Map position: 4
A;Introns: 225/1
A;Note: T16H5.90
C;Superfamily: Streptomyces chitinase chi40

  Query Match 67.7%; Score 44; DB 2; Length 332;
  Best Local Similarity 53.8%; Pred. No. 3.6;
  Matches 7; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VGYDDQESVKSKV 13
Db 320 IGYEDTQSVEAKV 332

RESULT 9
T04763
chitinase homolog T16H5.180 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-Jun-1999
C;Accession: T04763
R;Bevan, M.; De Haan, M.; Maarse, A.C.; Grivell, L.A.; Bancroft, I.; Mewes, H.W.; Mayer,
submitted to the Protein Sequence Database, June 1998
A;Reference number: Z15383
A;Accession: T04763
A;Molecule type: DNA
A;Residues: 1-366 <BEV>
A;Cross-references: EMBL:AL024486
A;Experimental source: cultivar Columbia; BAC clone T16H5
C;Genetics:
A;Map position: 4
A;Introns: 230/1
A;Note: T16H5.180
C;Superfamily: Streptomyces chitinase chi40

  Query Match 66.2%; Score 43; DB 2; Length 366;
  Best Local Similarity 53.8%; Pred. No. 6.1;
  Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VGYDDQESVKSKV 13
Db 320 IGYDDNQSIVMKV 332

RESULT 10
S51591
chitinase (EC 3.2.1.14) / lysozyme (EC 3.2.1.17) PZ precursor, pathogenesis-related
C;Species: Nicotiana tabacum (common tobacco)
C;Date: 15-Jul-1995 #sequence_revision 08-Sep-1995 #text_change 22-Jun-1999
C;Accession: S51591; S51632; S43119
R;Heitz, T.; Segond, S.; Kauffmann, S.; Geoffroy, P.; Prasad, V.; Brunner, F.; Frit
Mol. Gen. Genet. 245, 246-254, 1994
A;Title: Molecular characterization of a novel tobacco pathogenesis-related (PR) pr
A;Reference number: S51591; MUID:95115673; PMID:7816033
A;Accession: S51591
A;Molecule type: mRNA
A;Residues: 1-378 <HEI>
A;Cross-references: EMBL:X78325; NID:g467688; PIDN:CAA55128.1; PID:g467689
A;Experimental source: cv. Samsun NN
A;Accession: S51632
A;Molecule type: protein
A;Residues: 31-36;87-112;252-275;282-305;337-371 <HEW>
C;Superfamily: Streptomyces chitinase chi40
C;Keywords: glycosidase; hydrolase; polysaccharide degradation
F;1-25/Domain: signal sequence #status predicted <SIG>
F;26-371/Product: chitinase/lysozyme PZ #status experimental <MAT>
F;371-378/Domain: carboxyl-terminal propeptide #status predicted <PRO>

  Query Match 66.2%; Score 43; DB 2; Length 378;
  Best Local Similarity 53.8%; Pred. No. 6.3;
  Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 VGYDDQESVKSKV 13
Db 326 ISYDDTQSVRNKV 338

RESULT 11
T44445
chitinase (EC 3.2.1.14) [imported] - African malaria mosquito
C;Species: Anopheles gambiae (African malaria mosquito)
C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jan-2000
C;Accession: T44445
R;Shen, Z.; Jacobs-Lorena, M.
submitted to the EMBL Data Library, June 1997
A;Reference number: Z22771
A;Accession: T44445
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-525 <SHE>
A;Cross-references: EMBL:AF008575; PIDN:AAB87764.1
A;Experimental source: adult; gut
C;Genetics:
A;Gene: chi-1
C;Keywords: glycosidase; hydrolase

  Query Match 66.2%; Score 43; DB 2; Length 525;
  Best Local Similarity 69.2%; Pred. No. 8.9;
  Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VGYDDQESVKSKV 13
Db 347 VGYDDLRSVQLKV 359

RESULT 12
T04757
chitinase homolog T16H5.120 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-Jun-1999
C;Accession: T04757
R;Bevan, M.; De Haan, M.; Maarse, A.C.; Grivell, L.A.; Bancroft, I.; Mewes, H.W.; M
submitted to the Protein Sequence Database, June 1998

```
A;Reference number: Z15383
A;Accession: T04757
A;Molecule type: DNA
A;Residues: 1-365 <BEV>
A;Cross-references: EMBL:AL024486
A;Experimental source: cultivar Columbia; BAC clone T16H5
C;Genetics:
A;Map position: 4
A;Introns: 230/1
A;Note: T16H5.120
C;Superfamily: Streptomyces chitinase chi40

  Query Match            64.6%;  Score 42;  DB 2;  Length 365;
  Best Local Similarity  53.8%;  Pred. No. 9.1;
  Matches    7;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 VGYDDQESVKSKV 13
              :|||| :|:  |'
Db        321 IGYDDNQSIVYKV 333


RESULT 13
H81312
protein-export membrane protein Cj1093c [imported] - Campylobacter jejuni (strain NCTC 1
C;Species: Campylobacter jejuni
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 03-Jul-2002
C;Accession: H81312
R;Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chillin
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrel
Nature 403, 665-668, 2000
A;Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp
A;Reference number: A81250; MUID:20150912; PMID:10688204
A;Accession: H81312
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-525 <PAR>
A;Cross-references: GB:AL139077; GB:AL111168; NID:g6368444; PIDN:CAB73348.1; PID:g696852
A;Experimental source: serotype O2, strain NCTC 11168
C;Genetics:
A;Gene: secD; Cj1093c
C;Superfamily: protein export membrane protein secD

  Query Match            64.6%;  Score 42;  DB 2;  Length 526;
  Best Local Similarity  53.8%;  Pred. No. 14;
  Matches    7;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 VGYDDQESVKSKV 13
              :| '::|:|||.:
Db         51 LGVDNEEAVKSKI 63


RESULT 14
T24898
hypothetical protein T13H5.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T24898
R;Lightning, J.
submitted to the EMBL Data Library, October 1999
A;Reference number: Z19950
A;Accession: T24898
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-633 <WIL>
A;Cross-references: EMBL:Z66524; PIDN:CAA91419.1; GSPDB:GN00020; CESP:T13H5.3
A;Experimental source: clone T13H5
C;Genetics:
A;Gene: CESP:T13H5.3
A;Map position: 2
A;Introns: 27/1; 57/3; 93/3; 126/2; 251/2; 274/3; 457/3; 578/1

  Query Match            64.6%;  Score 42;  DB 2;  Length 633;
  Best Local Similarity  61.5%;  Pred. No. 16;

  Matches    8;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 VGYDDQESVKSKV 13
              |'|:: |||:| ||
Db        352 VGYENAESLKHKV 364


RESULT 15
G90165
conserved hypothetical protein [imported] - Sulfolobus solfataricus
C;Species: Sulfolobus solfataricus
C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 15-Jun-2001
C;Accession: G90165
R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; C
 Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redde
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A;Description: Sulfolobus solfataricus complete genome.
A;Reference number: A99139
A;Accession: G90165
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-349 <KUR>
A;Cross-references: GB:AE006641; NID:g13813378; PIDN:AAK40582.1; GSPDB:GN00155
C;Genetics:
A;Gene: SSO0241
C;Superfamily: conserved hypothetical protein MJ0135

  Query Match            61.5%;  Score 40;  DB 2;  Length 349;
  Best Local Similarity  61.5%;  Pred. No. 20;
  Matches    8;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 VGYDDQESVKSKV 13
              '||| | :|:||
Db         84 VGYDGIEKMKIKV 96


Search completed: September 24, 2003, 13:49:25
Job time : 16.7753 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 24, 2003, 13:26:37 ; Search time 8.1797# Seconds
(without alignments)
74.739 Million cell updates/sec

Title: US-09-744-282-7
Perfect score: 65
Sequence: 1 VGYDDQESVKSKV 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	65	100.0	383	1	C3L1_HUMAN	P36222 homo sapien
2	47	72.3	387	1	C3L1_MOUSE	Q61362 mus musculu
3	44	67.7	390	1	C3L2_HUMAN	Q15782 homo sapien
4	40	61.5	568	1	CAR7_CANAL	P43096 candida alb
5	38	58.5	499	1	GUN1_BACSU	P07983 bacillus su
6	38	58.5	499	1	GUN2_BACSU	P10475 bacillus su
7	37	56.9	400	1	GUN5_BACAG	O85465 bacillus aq
8	37	56.9	504	1	CHIT_BRUMA	P29030 brugia mala
9	37	56.9	554	1	CHI1_MANSE	P36362 manduca sex
10	37	56.9	725	1	HS9B_BRARE	O57521 brachydanio
11	37	56.9	970	1	AND1_MOUSE	P59328 mus musculu
12	36	55.4	116	1	LUXE_PHOLU	P19842 photorhabdu
13	36	55.4	266	1	VATD_NEUCR	O59941 neurospora
14	36	55.4	486	1	GATA_ANASP	Q8yy02 anabaena sp
15	36	55.4	499	1	CHIB_SERMA	P11797 serratia ma
16	36	55.4	575	1	YFBK_ECOLI	P76481 escherichia
17	36	55.4	797	1	DPOM_AGABI	P30322 agaricus bi
18	36	55.4	852	1	ACO2_HELPJ	Q9zi64 helicobacte
19	36	55.4	852	1	ACO2_HELPY	P56418 helicobacte
20	36	55.4	1704	1	CED7_CAERL	P34358 caenorhabdi
21	35	53.8	59	1	VXIS_BPL54	P20710 bacteriopha
22	35	53.8	124	1	VSP1_BPLMH	Q64765 lactococcus
23	35	53.8	145	1	MA29_DERFA	P39674 dermatophag
24	35	53.8	204	1	Y981_METJA	Q58390 methanococc
25	35	53.8	267	1	RL4_SULSO	Q9uxa6 sulfolobus
26	35	53.8	325	1	ISPH_CORGL	Q8nrm2 corynebacte
27	35	53.8	359	1	MURG_WIGBR	Q8d2z6 wiggleswort
28	35	53.8	374	1	RLIG_BPT4	P00971 bacteriopha
29	35	53.8	419	1	RHO_BUCAI	P57652 buchnera ap
30	35	53.8	419	1	RHO_BUCAP	O51891 buchnera ap
31	35	53.8	617	1	CHIT_CAEEL	Q11174 caenorhabdi
32	35	53.8	632	1	PABP_DROME	P21187 drosophila
33	35	53.8	699	1	CH11_BACCI	P20533 bacillus ci
34	35	53.8	850	1	PRSN_PIERA	Q9u8q4 pieris rapa
35	35	53.8	1553	1	TP2A_CHICK	O42130 gallus gall
36	35	53.8	4092	1	DYHC_YEAST	P36022 saccharomyc
37	34.5	53.1	642	1	DNAK_FRATU	P48205 francisella
38	34	52.3	277	1	KNOB_PLAFD	P05229 plasmodium
39	34	52.3	310	1	ARAC_ERWCH	P07642 erwinia chr
40	34	52.3	337	1	SYW_TREPA	O83640 treponema p
41	34	52.3	339	1	ANX2_XENLA	P24801 xenopus lae
42	34	52.3	339	1	ANXB_XENLA	P27006 xenopus lae
43	34	52.3	392	1	YHBH_BACSU	P45742 bacillus su
44	34	52.3	399	1	XFD1_XENLA	P33205 xenopus lae
45	34	52.3	441	1	MGTA_THEMA	P80099 thermotoga

ALIGNMENTS

```
RESULT 1
C3L1_HUMAN
ID   C3L1_HUMAN      STANDARD;      PRT;   383 AA.
AC   P36222; P30923;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Chitinase-3 like protein 1 precursor (Cartilage glycoprotein-39) (GP-
DE   39) (39 kDa synovial protein) (YKL-40).
GN   CHI3L1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC   TISSUE=Cartilage;
RX   MEDLINE=94064658; PubMed=8245017;
RA   Hakala B.E., White C., Recklies A.D.;
RT   "Human cartilage gp-39, a major secretory product of articular
RT   chondrocytes and synovial cells, is a mammalian member of a chitinase
RT   protein family.";
RL   J. Biol. Chem. 268:25803-25810(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Blood;
RX   MEDLINE=97386591; PubMed=9244440;
RA   Kehli M., Krause S.W., Andressen R.;
RT   "Molecular characterization of the gene for human cartilage gp-39
RT   (CHI3L1), a member of the chitinase protein family and marker for
RT   late stages of macrophage differentiation.";
RL   Genomics 43:221-225(1997).
RN   [3]
RP   SEQUENCE OF 22-45.
RX   MEDLINE=90328983; PubMed=2375755;
RA   Nyirkos P., Golds E.E.;
RT   "Human synovial cells secrete a 39 kDa protein similar to a bovine
RT   mammary protein expressed during the non-lactating period.";
RL   Biochem. J. 269:265-268(1990).
CC   -!- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN THE CAPACITY OF CELLS TO
CC       RESPOND TO AND COPE WITH CHANGES IN THEIR ENVIRONMENT.
CC   -!- SUBCELLULAR LOCATION: Extracellular.
CC   -!- TISSUE SPECIFICITY: PRESENT IN ARTICULAR CHONDROCYTES, SYNOVIAL
CC       CELLS AS WELL AS IN LIVER. UNDETECTABLE IN MUSCLE TISSUES, LUNG,
CC       PANCREAS, MONONUCLEAR CELLS, OR FIBROBLASTS.
CC   -!- PTM: GLYCOSYLATED.
CC   -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
```

```
DR   EMBL; M80927; AAA16074.1; -.
DR   EMBL; Y08374; CAA69661.1; -.
DR   EMBL; Y08375; CAA69661.1; JOINED.
DR   EMBL; Y08376; CAA69661.1; JOINED.
DR   EMBL; Y08377; CAA69661.1; JOINED.
DR   EMBL; Y08378; CAA69661.1; JOINED.
DR   PIR; A49562; A49562.
DR   PDB; 1LA7; 10-APR-02.
DR   Genew; HGNC:1932; CHI3L1.
DR   MIM; 601525; -.
DR   GO; GO:0005578; C:extracellular matrix; TAS.
DR   GO; GO:0005615; C:extracellular space; TAS.
DR   GO; GO:0005207; F:extracellular matrix glycoprotein; TAS.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   InterPro; IPR001579; Glyco_hydro_18/2.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   ProDom; PD000471; Glyco_hydro_18; 1.
DR   SMART; SM00636; Glyco_18; 1.
DR   PROSITE; PS01095; CHITINASE_18; FALSE_NEG.
KW   Glycoprotein; Signal; 3D-structure.
FT   SIGNAL        1     21
FT   CHAIN        22    383       CHITINASE-3 LIKE PROTEIN 1.
FT   CARBOHYD     60     60       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   383 AA;  42613 MW;  76ADD8298EEEC2D1 CRC64;

  Query Match             100.0%;  Score 65;  DB 1;  Length 383;
  Best Local Similarity   100.0%;  Pred. No. 0.00018;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VGYDDQESVKSKV 13
              |||||||||||||
Db        326 VGYDDQESVKSKV 338


RESULT 2
C3L1_MOUSE
ID   C3L1_MOUSE     STANDARD;      PRT:   381 AA.
AC   Q61362;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Chitinase-3 like protein 1 precursor (Cartilage glycoprotein-39)
DE   (GP-39) (BRP39 protein).
GN   CHI3L1 OR BRP39.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FVB/N; TISSUE=Breast;
RX   MEDLINE=95060797; PubMed=7973700;
RA   Morrison B.W., Leder P.;
RT   "neu and ras initiate murine mammary tumors that share genetic markers
RT   generally absent in c-myc and int-2-initiated tumors.";
RL   Oncogene 9:3417-3426(1994).
CC   -!- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN THE CAPACITY OF CELLS TO
CC       RESPOND TO AND COPE WITH CHANGES IN THEIR ENVIRONMENT (BY
CC       SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Extracellular (By similarity).
CC   -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.
CC   ---------------------------------------------------------------------- --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------- -
DR   EMBL; X93035; CAA63603.1; -.
DR   PIR; S61551; S61551.
DR   MGD; MGI:1340899; Chi3l1.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   InterPro; IPR001579; Glyco_hydro_18/2.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   ProDom; PD000471; Glyco_hydro_18; 1.
DR   SMART; SM00636; Glyco_18; 1.
DR   PROSITE; PS01095; CHITINASE_18; FALSE_NEG.
KW   Glycoprotein; Signal.
FT   SIGNAL        1     21       POTENTIAL.
FT   CHAIN        22    381       CHITINASE-3 LIKE PROTEIN 1.
FT   CARBOHYD     60     60       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   381 AA;  43001 MW;  EF6581E8184F0450 CRC64;

  Query Match             72.3%;  Score 47;  DB 1;  Length 381;
  Best Local Similarity   69.2%;  Pred. No. 0.39;
  Matches    9;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 VGYDDQESVKSKV 13
              ||| : ||||| :|
Db        327 VGYEHKESVKNKV 339


RESULT 3
C3L2_HUMAN
ID   C3L2_HUMAN     STANDARD;      PRT:   390 AA.
AC   Q15782; Q15749; Q15783;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Chitinase 3-like protein 2 precursor (YKL-39) (Chondrocyte protein
DE   39).
GN   CHI3L2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS LONG AND SHORT).
RA   Grossman A., Matsuyama T., Baker E., Waterhouse P., Sutherland G.R.,
RA   Mak T.W.;
RT   "Cloning of a novel lymphoid restricted human chitinase and
RT   localization to 1p13.3.";
RL   Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM SHORT).
RC   TISSUE=Articular cartilage;
RX   MEDLINE=96325055; PubMed=8702629;
RA   Hu B., Trinh K., Figueira W.F., Price P.A.;
RT   "Isolation and sequence of a novel human chondrocyte protein related
RT   to mammalian members of the chitinase protein family.";
RL   J. Biol. Chem. 271:19415-19420(1996).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=Short;
CC         IsoId=Q15782-1; Sequence=Displayed;
CC       Name=Long;
CC         IsoId=Q15782-2; Sequence=VSP_001066;
CC         Note=No experimental confirmation available;
CC   -!- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN CHONDROCYTES, FOLLOWED
CC       BY SYNOVIOCYTES, LUNG AND HEART. NOT DETECTED IN BRAIN, SPLEEN,
CC       PANCREAS, AND LIVER.
CC   -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.
CC   ---------------------------------------------------------------------- --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------- --
DR   EMBL; U58515; AAB04534.1; -.
DR   EMBL; U58514; AAB04533.1; -.
DR   EMBL; U49835; AAC50597.1; ALT_INIT.
```

```
DR   Genew; HGNC:1933; CHI3L2.
DR   MIM; 601526; -.
DR   GO; GO:0005615; C:extracellular space; TAS.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   InterPro; IPR001579; Glyco_hydro_18/2.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   ProDom; PD000471; Glyco_hydro_18; 1.
DR   SMART; SM00636; Glyco_18; 1.
DR   PROSITE; PS01095; CHITINASE_18; FALSE_NEG.
KW   Hydrolase; Signal; Alternative splicing.
FT   SIGNAL       1     26       POTENTIAL.
FT   CHAIN       27    390       CHITINASE 3-LIKE PROTEIN 2.
FT   CARBOHYD    35     35       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC     1     12       MGATTMDQKSLW -> SQQLWLGSPDEVWLYLVCEHTHIFT
FT                               AIIWDSRTRPGSTDFSPHPSI (in isoform Long).
FT                               /FTId=VSP_001066.
SQ   SEQUENCE   390 AA;  43500 MW;  97B86A2F3AA35677 CRC64;

  Query Match             67.7%;  Score 44;  DB 1;  Length 390;
  Best Local Similarity   61.5%;  Pred. No. 1.4;
  Matches    8;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 VGYDDQFSVKSKV 13
              :|||. :|:::|:
Db        334 VGYDDVKSMEIKV 346


RESULT 4
CAR7_CANAL
ID   CAR7_CANAL     STANDARD;      PRT:   586 AA.
AC   P43096;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Candidapepsin 7 precursor (EC 3.4.23.24) (Aspartate protease 7)
DE   (ACP 7) (Secreted aspartic protease 7).
GN   SAP7.
OS   Candida albicans (Yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX   NCBI_TaxID=5476;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C74;
RX   MEDLINE=95075329; PubMed=7984113;
RA   Monod M., Togni G., Hube B., Sanglard D.;
RT   "Multiplicity of genes encoding secreted aspartic proteinases in
RT   Candida species.";
RL   Mol. Microbiol. 13:357-368(1994).
CC   -!- CATALYTIC ACTIVITY: Preferential cleavage at the carboxyl of
CC       hydrophobic amino acids, but fails to cleave 15-Leu-|-Tyr-16, 16-
CC       Tyr-|-Leu-17 and 24-Phe-|-Phe-25 of insulin B chain. Activates
CC       trypsinogen, and degrades keratin.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- PTM: O-GLYCOSYLATED (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z30193; CAA82925.1; -.
DR   HSSP; P32329; 1YPS.
DR   MEROPS; A01.UPW; -.
DR   InterPro; IPR001969; Aspprotease_site.
DR   InterPro; IPR001461; AspproteaseA1.
DR   Pfam; PF00026; asp; 1.
DR   PROSITE; PS00141; ASP_PROTEASE; 1.
KW   Hydrolase; Aspartyl protease; Glycoprotein; Zymogen; Signal;
KW   Multigene family.
FT   SIGNAL       1     16       POTENTIAL.
FT   PROPEP      17    211       POTENTIAL.
FT   CHAIN      212    588       CANDIDAPEPSIN 7.
FT   ACT_SITE   244    244       BY SIMILARITY.
FT   ACT_SITE   464    464       BY SIMILARITY.
FT   DISULFID   260    269       BY SIMILARITY.
FT   DISULFID   500    540       BY SIMILARITY.
FT   CARBOHYD   150    150       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   286    286       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   308    308       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   327    327       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   423    423       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   445    445       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   486    486       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   DOMAIN     180    190       POLY-SER.
SQ   SEQUENCE   588 AA;  62553 MW;  01C676FC39C21D61 CRC64;

  Query Match             61.5%;  Score 40;  DB 1;  Length 588;
  Best Local Similarity   66.7%;  Pred. No. 14;
  Matches    8;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 VGYDDQFSVKSK 12
              ||| : |:|||||
Db        337 VGYPELEAVKSK 348


RESULT 5
GUN1_BACSU
ID   GUN1_BACSU     STANDARD;      PRT:   499 AA.
AC   P07983;
DT   01-AUG-1988 (Rel. 08, Created)
DT   01-NOV-1991 (Rel. 20, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Endoglucanase precursor (EC 3.2.1.4) (Endo-1,4-beta-glucanase)
DE   (Cellulase).
GN   BGLC OR CLD.
OS   Bacillus subtilis.
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX   NCBI_TaxID=1423;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DLG;
RX   MEDLINE=87194581; PubMed=3106328;
RA   Robson L.M., Chambliss G.H.;
RT   "Endo-beta-1,4-glucanase gene of Bacillus subtilis DLG.";
RL   J. Bacteriol. 169:2017-2025(1987).
CC   -!- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic
CC       linkages in cellulose, lichenin and cereal beta-D-glucans.
CC   -!- SIMILARITY: BELONGS TO CELLULASE FAMILY A (FAMILY 5 OF GLYCOSYL
CC       HYDROLASES).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M16185; AAA22496.1; ALT_INIT.
DR   HSSP; O85465; 1A3H.
DR   InterPro; IPR001956; CBD_3.
DR   InterPro; IPR001547; Glyco_hydro_5.
DR   Pfam; PF00942; CBM_3; 1.
DR   Pfam; PF00150; cellulase; 1.
DR   ProDom; PD001947; CBD_3; 1.
DR   PROSITE; PS00659; GLYCOSYL_HYDROL_F5; 1.
KW   Cellulose degradation; Hydrolase; Glycosidase; Signal.
FT   SIGNAL       1     29
FT   CHAIN       30    499       ENDOGLUCANASE.
FT   ACT_SITE   169    169       PROTON DONOR (BY SIMILARITY).
FT   ACT_SITE   257    257       NUCLEOPHILE (BY SIMILARITY).
```

```
FT   DOMAIN       350    499       CELLULOSE BINDING (BY SIMILARITY).
SQ   SEQUENCE   499 AA;  55187 MW;  339D04BE95A63EF1 CRC64;

  Query Match             58.5%;  Score 38;  DB 1;  Length 499;
  Best Local Similarity   66.7%;  Pred. No. 24;
  Matches    8;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 GYDDQESVKSKV 13
              || |  |||:||
Db        101 GYIDNPSVKNKV 112


RESULT 6
GUN2_BACSU
ID   GUN2_BACSU     STANDARD;      PRT;   499 AA.
AC   P10475;
DT   01-JUL-1988 (Rel. 11, Created)
DT   01-JUL-1988 (Rel. 11, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Endoglucanase precursor (EC 3.2.1.4) (Endo-1,4-beta-glucanase)
DE   (Carboxymethyl-cellulase) (CMCASE) (Cellulase).
GN   BGLC OR GLD OR EGLS.
OS   Bacillus subtilis.
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX   NCBI_TaxID=1423;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PAP115;
RX   MEDLINE=87066783; PubMed=3024130;
RA   Mackay R.M., Lo A., Willick G., Zuker M., Baird S., Dove M.,
RA   Moranelli F., Seligy V.;
RT   "Structure of a Bacillus subtilis endo-beta-1,4-glucanase gene.";
RL   Nucleic Acids Res. 14:9159-9170(1986).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CK-2;
RX   MEDLINE=95225656; PubMed=7710280;
RA   Lindahl V., Aa K., Tronsmo A.;
RT   "Nucleotide sequence of an endo-beta-1,4-glucanase gene from Bacillus
RT   subtilis CK-2.";
RL   Antonie Van Leeuwenhoek 66:327-332(1994).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168;
RX   MEDLINE=97124194; PubMed=8969507;
RA   Rose M., Entian K.D.;
RT   "New genes in the 170 degrees region of the Bacillus subtilis genome
RT   encode DNA gyrase subunits, a thioredoxin, a xylanase and an amino
RT   acid transporter.";
RL   Microbiology 142:3097-3101(1996).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168;
RX   MEDLINE=98044033; PubMed=9384377;
RA   Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA   Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA   Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA   Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA   Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA   Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA   Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA   Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA   Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA   Guiseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
RA   Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA   Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA   Kobayashi Y., Koetter P., Koningstein G., Krogh S., Kumano M.,
RA   Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA   Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA   Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA   Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA   Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA   Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA   Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA   Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA   Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA   Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA   Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA   Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA   Viari A., Wambutt R., Wedler E., Wedler H., Weitzenegger T.,
RA   Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA   Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RT   "The complete genome sequence of the Gram-positive bacterium Bacillus
RT   subtilis.";
RL   Nature 390:249-256(1997).
RN   [5]
RP   SEQUENCE OF 30-45.
RC   STRAIN=CK-2;
RX   MEDLINE=95225655; PubMed=7710279;
RA   Aa K., Flengsrud R., Lindahl V., Tronsmo A.;
RT   "Characterization of production and enzyme properties of an
RT   endo-beta-1,4-glucanase from Bacillus subtilis CK-2 isolated from
RT   compost soil.";
RL   Antonie Van Leeuwenhoek 66:319-326(1994).
CC   -!- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic
CC       linkages in cellulose, lichenin and cereal beta-D-glucans.
CC   -!- SIMILARITY: BELONGS TO CELLULASE FAMILY A (FAMILY 5 OF GLYCOSYL
CC       HYDROLASES).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z29076; CAA82317.1; -.
DR   EMBL; X04689; CAA28392.1; -.
DR   EMBL; X67044; CAA47429.1; -.
DR   EMBL; Z73234; CAA97610.1; ALT_INIT.
DR   EMBL; Z99113; CAB13696.1; ALT_INIT.
DR   PIR; G69593; G69593.
DR   HSSP; O85465; 1A3H.
DR   SubtiList; BG10437; bglC.
DR   InterPro; IPR001956; CBD_3.
DR   InterPro; IPR001547; Glyco_hydro_5.
DR   Pfam; PF00942; CBM_3; 1.
DR   Pfam; PF00150; cellulase; 1.
DR   ProDom; PD001947; CBD_3; 1.
DR   PROSITE; PS00659; GLYCOSYL_HYDROL_F5; 1.
KW   Cellulose degradation; Hydrolase; Glycosidase; Signal;
KW   Complete proteome.
FT   SIGNAL       1     29
FT   CHAIN       30    499       ENDOGLUCANASE.
FT   ACT_SITE    169    169       PROTON DONOR (BY SIMILARITY).
FT   ACT_SITE    257    257       NUCLEOPHILE (BY SIMILARITY).
FT   DOMAIN      350    499       CELLULOSE-BINDING (BY SIMILARITY).
FT   CONFLICT    283    283       S -> N (IN REF. 2).
SQ   SEQUENCE   499 AA;  55287 MW;  8F735FF711B3EA22 CRC64;

  Query Match             58.5%;  Score 38;  DB 1;  Length 499;
  Best Local Similarity   66.7%;  Pred. No. 24;
  Matches    8;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 GYDDQESVKSKV 13
              || |  |||:||
Db        101 GYIDNPSVKNKV 112


RESULT 7
GUN5_BACAG
ID   GUN5_BACAG     STANDARD;      PRT;   400 AA.
AC   O85465;
DT   15-JUL-1999 (Rel. 38, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
```

```
DI   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Endoglucanase 5A (EC 3.2.1.4) (Endo-1,4-beta-glucanase) (Alkaline
DE   cellulase).
GN   CEL5A.
OS   Bacillus agaradhaerens (Bacillus agaradherans).
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX   NCBI_TaxID=76935;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DSM 8721;
RA   Bjornvad M.E.;
RL   Submitted (MAY-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   X-RAY CRYSTALLOGRAPHY (1.57 ANGSTROMS) OF 30-329.
RC   STRAIN=AC13 / NCIMB 40482;
RX   MEDLINE=98153671; PubMed=9485319;
RA   Davies G.J., Dauter M., Brzozowski A.M., Bjoernvad M.E.,
RA   Andersen K.V., Schuelein M.;
RT   "Structure of the Bacillus agaradherans family 5 endoglucanase at 1.6-
RT   A and its cellobiose complex at 2.0-A resolution.";
RL   Biochemistry 37:1926-1932(1998).
RN   [3]
RP   X-RAY CRYSTALLOGRAPHY (1.64 ANGSTROMS) OF 30-329.
RC   STRAIN=AC13 / NCIMB 40482;
RX   MEDLINE=98384136; PubMed=9718293;
RA   Davies G.J., MacKenzie L.F., Varrot A., Dauter M., Brzozowski A.M.,
RA   Schuelein M., Withers S.G.;
RT   "Snapshots along an enzymatic reaction coordinate: analysis of a
RT   retaining beta-glycoside hydrolase.";
RL   Biochemistry 37:11707-11713(1998).
CC   -!- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic
CC       linkages in cellulose, lichenin and cereal beta-D glucans.
CC   -!- SUBUNIT: Monomer.
CC   -!- SIMILARITY: BELONGS TO CELLULASE FAMILY A (FAMILY 5 OF GLYCOSYL
CC       HYDROLASES).
CC   ---------------------------------------------------------- ----------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------- - --------
DR   EMBL; AF067428; AAC19169.1; -.
DR   PDB; 1A3H; 16-MAR-99.
DR   PDB; 2A3H; 16-MAR-99.
DR   PDB; 3A3H; 16-MAR-99.
DR   PDB; 1H11; 27-FEB-03.
DR   PDB; 1HF6; 29-NOV-01.
DR   PDB; 5A3H; 24-JUL-98.
DR   PDB; 6A3H; 24-JUL-98.
DR   PDB; 7A3H; 06-AUG-99.
DR   InterPro; IPR003610; CBM_5_12.
DR   InterPro; IPR001547; Glyco_hydro_5.
DR   Pfam; PF02839; CBM_5_12; 1.
DR   Pfam; PF00150; cellulase; 1.
DR   SMART; SM00495; ChtBD3; 1.
DR   PROSITE; PS00659; GLYCOSYL_HYDROL_F5; 1.
KW   Cellulose degradation; Hydrolase; Glycosidase; 3D-structure.
FT   ACT_SITE    165    165       PROTON DONOR.
FT   ACT_SITE    254    254       NUCLEOPHILE.
FT   HELIX        31     35
FT   STRAND       39     41
FT   TURN         42     43
FT   STRAND       44     46
FT   TURN         48     49
FT   STRAND       52     52
FT   STRAND       56     60
FT   HELIX        63     66
FT   HELIX        67     69
FT   HELIX        72     81
FT   TURN         82     82
FT   STRAND       86     92
FT   TURN         95     96
FT   TURN         98    100
FT   TURN        102    103
FT   HELIX       104    118
FT   TURN        119    119
FT   STRAND      121    127
FT   TURN        134    137
FT   HELIX       138    152
FT   TURN        153    154
FT   TURN        156    157
FT   STRAND      158    161
FT   TURN        169    170
FT   TURN        173    176
FT   HELIX       177    189
FT   TURN        190    191
FT   STRAND      197    199
FT   HELIX       202    205
FT   TURN        206    206
FT   STRAND      207    207
FT   HELIX       208    212
FT   TURN        213    213
FT   TURN        219    220
FT   STRAND      221    228
FT   TURN        229    230
FT   HELIX       234    245
FT   TURN        246    247
FT   STRAND      250    256
FT   TURN        260    261
FT   HELIX       268    280
FT   TURN        281    282
FT   STRAND      285    290
FT   TURN        296    297
FT   STRAND      300    300
FT   TURN        302    303
FT   TURN        306    307
FT   HELIX       312    314
FT   STRAND      315    315
FT   HELIX       317    328
SQ   SEQUENCE   400 AA;  44702 MW;  3F9C65FB9BC35FFF CRC64;

  Query Match           56.9%;  Score 37;  DF 1;  Length 400;
  Best Local Similarity  66.7%;  Pred. No. 29;
  Matches    8;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 GYDDQESVKSKV 13
              || |  ||| |:
Db         97 GYIDDPSVKEKV 108


RESULT 6
CHIT_BRUMA
ID   CHIT_BRUMA     STANDARD;      PRT;   504 AA.
AC   P29030;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Endochitinase precursor (EC 3.2.1.14) (MF1 antigen).
OS   Brugia malayi (Filarial nematode worm).
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
OC   Onchocercidae; Brugia.
OX   NCBI_TaxID=6279;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX   MEDLINE=92179220; PubMed=1542646;
RA   Fuhrman J.A., Lane W.S., Smith R.F., Piessens W.F., Perler F.B.;
RT   "Transmission-blocking antibodies recognize microfilarial chitinase
RT   in brugian lymphatic filariasis.";
RL   Proc. Natl. Acad. Sci. U.S.A. 89:1548-1552(1992).
CC   -!- FUNCTION: THE MF1 ANTIGEN IS A MICROFILARIAL CHITINASE, WHICH MAY
CC       FUNCTION TO DEGRADE CHITIN-CONTAINING STRUCTURES IN THE MICRO-
CC       FILARIA OR IN ITS MOSQUITO VECTOR DURING PARASITE DEVELOPMENT AND
CC       TRANSMISSION.
```

```
CC   -!- CATALYTIC ACTIVITY: Hydrolysis of the 1,4-beta-linkages of N-
CC       acetyl-D-glucosamine polymers of chitin.
CC   -!- DEVELOPMENTAL STAGE: THE APPEARANCE OF THE MF1 ANTIGEN CORRESPOND
CC       WITH THE ONSET OF THE PARASITE'S ABILITY TO INFECT THE MOSQUITO.
CC   -!- PTM: O-GLYCOSYLATED.
CC   -!- MISCELLANEOUS: KNOWN TO BIND CALCIUM.
CC   -!- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
CC       HYDROLASES).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M73689; AAA27854.1; -.
DR   PIR; A38221; A38221.
DR   InterPro; IPR002557; Chitin_bind_PerA.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   InterPro; IPR001579; Glyco_hydro_18/2.
DR   Pfam; PF01607; CBM_14; 1.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   ProDom; PD000471; Glyco_hydro_18; 1.
DR   SMART; SM00494; ChtBD2; 1.
DR   SMART; SM00636; Glyco_18; 1.
DR   PROSITE; PS01095; CHITINASE_18; 1.
KW   Hydrolase; Glycosidase; Chitin degradation; Signal; Antigen;
KW   Repeat; Glycoprotein; Calcium-binding.
FT   SIGNAL        1     22
FT   CHAIN        23    504       ENDOCHITINASE.
FT   DOMAIN       23    400       CATALYTIC.
FT   DOMAIN      401    450       SER/THR-RICH (LINKER).
FT   DOMAIN      407    448       3 X 14 AA APPROXIMATE TANDEM REPEATS.
FT   ACT_SITE    148    148       PROTON DONOR (BY SIMILARITY).
SQ   SEQUENCE   504 AA;  55971 MW;  A78BE7BFB8E37095 CRC64;

  Query Match             56.9%;  Score 37;  DB 1;  Length 504;
  Best Local Similarity   41.7%;  Pred. No. 38;
  Matches    5;  Conservative    6;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 GYDDQESVKSKV 13
              |||:|:::: |:
Db        337 GYDNEETIRIKM 348


RESULT 9
CHIT_MANSE
ID   CHIT_MANSE     STANDARD;      PRT;   554 AA.
AC   P36362;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-FEB-1995 (Rel. 31, Last annotation update)
DE   Endochitinase precursor (EC 3.2.1.14).
OS   Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Sphingoidea;
OC   Sphingidae; Sphinginae; Manduca.
OX   NCBI_TaxID=7130;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93357793; PubMed=8353525;
RA   Kramer K.J., Corpuz L., Choi H.K., Muthukrishnan S.;
RT   "Sequence of a cDNA and expression of the gene encoding epidermal and
RT   gut chitinases of Manduca sexta.";
RL   Insect Biochem. Mol. Biol. 23:691-701(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97215580; PubMed=9061927;
RA   Choi H.K., Choi K.H., Kramer K.J., Muthukrishnan S.;
RT   "Isolation and characterization of a genomic clone for the gene of an
RT   insect molting enzyme, chitinase.";
RL   Insect Biochem. Mol. Biol. 27:37-47(1997).
CC   -!- FUNCTION: DIGEST CHITIN IN THE EXOSKELETON DURING THE MOLTING
CC       PROCESS.
CC   -!- CATALYTIC ACTIVITY: Hydrolysis of the 1,4-beta-linkages of N-
CC       acetyl-D-glucosamine polymers of chitin.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: EPIDERMIS AND GUT.
CC   -!- DEVELOPMENTAL STAGE: HIGH LEVELS SEEN IN THE EPIDERMIS ON DAY 0,
CC       BUT RAPIDLY DISAPPEARS AND IS UNDETECTED ON DAYS 1-4 OF FIFTH
CC       INSTAR. IT REAPPEARS ON DAY 5 AND PEAKS ON DAY 7 AFTER WHICH A
CC       RAPID DECLINE IS SEEN. IN THE GUT IS DETECTED ON DAY 6 WITH LOWER
CC       LEVELS SEEN ON DAYS 0, 7 AND 8.
CC   -!- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
CC       HYDROLASES).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U02270; AAC04924.1; -.
DR   EMBL; L49234; AAB53952.1; -.
DR   PIR; A56596; A56596.
DR   InterPro; IPR002557; Chitin_bind_PerA.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   InterPro; IPR001579; Glyco_hydro_18/2.
DR   Pfam; PF01607; CBM_14; 1.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   ProDom; PD000471; Glyco_hydro_18; 1.
DR   SMART; SM00494; ChtBD2; 1.
DR   SMART; SM00636; Glyco_18; 1.
DR   PROSITE; PS01095; CHITINASE_18; 1.
KW   Hydrolase; Glycosidase; Chitin degradation; Signal; Glycoprotein.
FT   SIGNAL        1     19       POTENTIAL.
FT   CHAIN        20    554       ENDOCHITINASE.
FT   DOMAIN      396    453       SER/THR-RICH.
FT   ACT_SITE    146    146       PROTON DONOR (BY SIMILARITY).
FT   CARBOHYD     85     85       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    303    303       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    407    407       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    545    545       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   554 AA;  62203 MW;  3989D756C96CD490 CRC64;

  Query Match             56.9%;  Score 37;  DB 1;  Length 554;
  Best Local Similarity   53.8%;  Pred. No. 42;
  Matches    7;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 VGYDDQESVKSKV 13
              |||:|  ||: |:
Db        344 VGYEDPRSVEIKM 356


RESULT 10
HS9B_BRARE
ID   HS9B_BRARE     STANDARD;      PRT;   725 AA.
AC   O57521; Q90475;
DT   15-JUL-1999 (Rel. 38, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Heat shock protein HSP 90-beta.
GN   HSP90B.
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Coumailleau P., Angelier N.;
RT   "Molecular cloning of a cDNA encoding the Danio rerio hsp90 beta
```

```
RT   isoform.";
RL   Submitted (JAN-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Lele Y., Hadfi S., Sass J.B., Krone P.H.;
RL   Submitted (MAY-1998) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 32-132 FROM N.A.
RX   MEDLINE=95071389; PubMed=7980538;
RA   Krone P.H., Sass J.B.;
RT   "HSP 90 alpha and HSP 90 beta genes are present in the zebrafish and
RT   are differentially regulated in developing embryos.";
RL   Biochem. Biophys. Res. Commun. 204:746-752(1994).
CC   -!- FUNCTION: MOLECULAR CHAPERONE. HAS ATPASE ACTIVITY
CC       (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 90 FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF042198; AAB96969.1; -.
DR   EMBL; AF068772; AAC21566.1; -.
DR   EMBL; L35587; AAA97519.1; -.
DR   HSSP; P07900; 1YER.
DR   ZFIN; ZDB-GENE-990415-95; hsp90b.
DR   InterPro; IPR003594; ATPbind_ATPase.
DR   InterPro; IPR001404; Hsp90.
DR   Pfam; PF02518; HATPase_c; 1.
DR   Pfam; PF00183; HSP90; 1.
DR   PRINTS; PR00775; HEATSHOCK90.
DR   SMART; SM00387; HATPase_c; 1.
DR   PROSITE; PS00298; HSP90; 1.
KW   Chaperone; ATP-binding; Heat shock; Phosphorylation.
FT   MOD_RES     225     225       PHOSPHORYLATION (BY SIMILARITY).
FT   MOD_RES     254     254       PHOSPHORYLATION (BY SIMILARITY).
FT   CONFLICT     98      98       I -> M (IN REF. 3).
FT   CONFLICT    417     418       EL -> DV (IN REF. 2).
FT   CONFLICT    444     444       C -> S (IN REF. 2).
FT   CONFLICT    448     448       K -> R (IN REF. 2).
FT   CONFLICT    463     463       D -> Y (IN REF. 2).
FT   CONFLICT    622     622       N -> K (IN REF. 2).
FT   CONFLICT    647     647       T -> D (IN REF. 2).
FT   CONFLICT    701     703       TSA -> SS (IN REF. 2).
SQ   SEQUENCE   725 AA;  83359 MW;  57D612F33773A8C1 CRC64;

  Query Match             56.9%;  Score 37;  DB 1;  Length 725;
  Best Local Similarity   58.3%;  Pred. No. 57;
  Matches    7;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 VGYDDQESVKSK 12
              || |:|  | |
Db        252 VGSDDREDIKDK 263


RESULT 11
AND1_MOUSE
ID   AND1_MOUSE     STANDARD;      PRT;   970 AA.
AC   P59328;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Acidic nucleoplasmic DNA-binding protein 1 (And-1) (Fragment).
GN   AND1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Bone, and Kidney;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
CC   -!- FUNCTION: Binds DNA with high affinity. May also be involved in
CC       protein-protein interactions (By similarity).
CC   -!- SUBCELLULAR LOCATION: Nuclear; nucleoplasm (By similarity).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1;
CC         IsoId=P59328-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=P59328-2; Sequence=VSP_006757, VSP_006758;
CC         Note=No experimental confirmation available;
CC   -!- SIMILARITY: Contains 7 WD repeats.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AK036390; BAC29408.1; -.
DR   EMBL; AK052690; BAC35097.1; -.
DR   MGD; MGI:2443514; D630024B06Rik.
DR   Pfam; PF00400; WD40; 4.
DR   ProDom; PD000018; WD40; 1.
DR   PROSITE; PS00678; WD_REPEATS_1; 2.
DR   PROSITE; PS50082; WD_REPEATS_2; 3.
DR   PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW   Nuclear protein; DNA-binding; Repeat; WD repeat; Alternative splicing.
FT   REPEAT       11      50       WD 1.
FT   REPEAT       52      91       WD 2.
FT   REPEAT       93     131       WD 3.
FT   REPEAT      134     173       WD 4.
FT   REPEAT      184     223       WD 5.
FT   REPEAT      228     267       WD 6.
FT   REPEAT      271     310       WD 7.
FT   VARSPLIC    630     641       TPCYVDSEGCVR -> KIIFFLYIKDIF (in isoform
FT                                 2).
FT                                 /FTId=VSP_006757.
FT   VARSPLIC    642    >970       Missing (in isoform 2).
FT                                 /FTId=VSP_006758.
FT   CONFLICT     54      54       G -> V (IN REF. 1; BAC29408).
FT   CONFLICT    350     350       S -> P (IN REF. 1; BAC29408).
FT   NON_TER     970     970
SQ   SEQUENCE   970 AA;  107632 MW;  2B2B577A9A6D2792 CRC64;

  Query Match             56.9%;  Score 37;  DB 1;  Length 970;
  Best Local Similarity   54.5%;  Pred. No. 79;
  Matches    6;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;
```

```
Qy          2 GYDDQESVKSK 12
              ||| :|||:|::
Db        739 GYDYEESIKNQ 749


RESULT 12
LUXE_PHOLU
ID   LUXE_PHOLU     STANDARD;      PRT;   116 AA.
AC   P19842;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-NOV-1991 (Rel. 20, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   Long-chain-fatty-acid--luciferin-component ligase (EC 6.2.1.18) (Acyl-
DE   protein synthetase) (Fragment).
GN   LUXE.
OS   Photorhabdus luminescens (Xenorhabdus luminescens).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Photorhabdus.
OX   NCBI_TaxID=29488;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HW;
RX   MEDLINE=91139581; PubMed=1995589;
RA   Xi L., Cho K.W., Tu S.C.;
RT   "Cloning and nucleotide sequences of lux genes and characterization
RT   of luciferase of Xenorhabdus luminescens from a human wound.";
RL   J. Bacteriol. 173:1399-1405(1991).
RN   [2]
RP   SEQUENCE OF 1-37 FROM N.A.
RC   STRAIN=HM;
RX   MEDLINE=90343746; PubMed=2383248;
RA   Johnston T.C., Rucker E.B., Cochrum D., Hruska K.S., Vandegrift V.;
RT   "The nucleotide sequence of the luxA and luxB genes of Xenorhabdus
RT   luminescens HM and a comparison of the amino acid sequences of
RT   luciferases from four species of bioluminescent bacteria.";
RL   Biochem. Biophys. Res. Commun. 170:407-415(1990).
RN   [3]
RP   SEQUENCE OF 1-4 FROM N.A.
RC   STRAIN=ATCC 29999;
RX   MEDLINE=90375532; PubMed=2204626;
RA   Szittner R., Meighen E.;
RT   "Nucleotide sequence, expression, and properties of luciferase coded
RT   by lux genes from a terrestrial bacterium.";
RL   J. Biol. Chem. 265:16581-16587(1990).
CC   -!- FUNCTION: ACYL-PROTEIN SYNTHETASE ACTIVATES TETRADECANOIC ACID.
CC       IT IS A COMPONENT OF THE FATTY ACID REDUCTASE COMPLEX RESPONSIBLE
CC       FOR CONVERTING TETRADECANOIC ACID TO THE ALDEHYDE WHICH SERVES AS
CC       SUBSTRATE IN THE LUCIFERASE-CATALYZED REACTION.
CC   -!- CATALYTIC ACTIVITY: ATP + an acid + protein = AMP + diphosphate +
CC       an acyl-protein thiolester.
CC   -!- PATHWAY: Bioluminescent fatty acid reduction system; second step.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M62917; AAA63567.1; -.
DR   EMBL; M57416; -; NOT_ANNOTATED_CDS.
DR   EMBL; M55977; AAA27626.1; -.
DR   Pfam; PF04443; LuxE; 1.
KW   Luminescence; Ligase.
FT   NON_TER     116    116
SQ   SEQUENCE   116 AA;  13575 MW;  4E7C717FD749CAC5 CRC64;

  Query Match             55.4%;  Score 36;  DB 1;  Length 116;
  Best Local Similarity   45.5%;  Pred. No. 11;
  Matches    5;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          3 YDDQESVKSKV 13
              ||:|| :: |:
Db         30 YDEQEKIRKKL 40


RESULT 13
VATD_NEUCR
ID   VATD_NEUCR     STANDARD;      PRT;   266 AA.
AC   O59941;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Vacuolar ATP synthase subunit D (EC 3.6.3.14) (V-ATPase D subunit)
DE   (Vacuolar proton pump D subunit).
GN   VMA-8.
OS   Neurospora crassa.
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC   Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX   NCBI_TaxID=5141;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99270694; PubMed=10340846;
RA   Margolles-Clark E., Tenney K., Bowman E.J., Bowman B.J.;
RT   "The structure of the vacuolar ATPase in Neurospora crassa.";
RL   J. Bioenerg. Biomembr. 31:29-37(1999).
CC   -!- FUNCTION: SUBUNIT OF THE PERIPHERAL V1 COMPLEX OF VACUOLAR ATPASE.
CC       V-ATPASE IS RESPONSIBLE FOR ACIDIFYING A VARIETY OF INTRACELLULAR
CC       COMPARTMENTS IN EUKARYOTIC CELLS. THUS PROVIDING MOST OF THE
CC       ENERGY REQUIRED FOR TRANSPORT PROCESSES IN THE VACUOLAR SYSTEM.
CC   -!- CATALYTIC ACTIVITY: ATP + H(2)O + H(+)(In) = ADP + phosphate +
CC       H(+)(Out).
CC   -!- SUBUNIT: V-ATPase is an heteromultimeric enzyme composed of a
CC       peripheral catalytic V1 complex (components A to H) attached to
CC       an integral membrane V0 proton pore complex (components: a, c, c',
CC       c'' and d).
CC   -!- SIMILARITY: Belongs to the V-ATPase D subunit family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AF053230; AAC08354.1; -.
DR   InterPro; IPR002699; ATPsynt_Dsub.
DR   Pfam; PF01813; ATP-synt_D; 1.
DR   ProDom; PD004122; ATPsynt_Dsub; 1.
DR   TIGRFAMs; TIGR00309; V_ATPase_subD; 1.
KW   Hydrolase; ATP synthesis; Hydrogen ion transport.
SQ   SEQUENCE   266 AA;  29735 MW;  361FAB084BD84904 CRC64;

  Query Match             55.4%;  Score 36;  DB 1;  Length 266;
  Best Local Similarity   63.6%;  Pred. No. 28;
  Matches    7;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 VGYDDQESVKS 11
              :|| |:| ||
Db         81 IGYQIQESAKS 91


RESULT 14
GATA_ANASP
ID   GATA_ANASP     STANDARD;      PRT;   486 AA.
AC   Q8YY02;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Glutamyl-tRNA(Gln) amidotransferase subunit A (EC 6.3.5.-) (Glu-ADT
DE   subunit A).
GN   GATA OR ALL1053.
OS   Anabaena sp. (strain PCC 7120).
```

```
OC   Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX   NCBI_TaxID=103690;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21595285; PubMed=11759840;
RA   Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA   Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA   Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA   Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,
RA   Yasuda M., Tabata S.;
RT   "Complete genomic sequence of the filamentous nitrogen-fixing
RT   cyanobacterium Anabaena sp. strain PCC 7120.";
RL   DNA Res. 8:205-213(2001).
CC   -!- FUNCTION: Furnishes a means for formation of correctly charged
CC       Gln-tRNA(Gln) through the transamidation of misacylated Glu-
CC       tRNA(Gln) in organisms which lack glutaminyl-tRNA synthetase. The
CC       reaction takes place in the presence of glutamine and ATP through
CC       an activated gamma-phospho-Glu-tRNA(Gln) (By similarity).
CC   -!- CATALYTIC ACTIVITY: ATP + L-glutamyl-tRNA(Gln) + L-glutamine = ADP
CC       + phosphate + L-glutaminyl-tRNA(Gln) + L-glutamate.
CC   -!- SUBUNIT: Heterotrimer of A, B and C subunits (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE AMIDASE FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AP003584; BAB73010.1; -.
DR   PIR; AB1938; AB1938.
DR   HAMAP; MF_00120; -; 1.
DR   InterPro; IPR000120; Amidase.
DR   InterPro; IPR004412; GatA.
DR   Pfam; PF01425; Amidase; 1.
DR   TIGRFAMs; TIGR00132; gatA; 1.
DR   PROSITE; PS00571; AMIDASES; 1.
KW   Protein biosynthesis; Ligase; Complete proteome.
SQ   SEQUENCE   486 AA;  52373 MW;  C3BA84B3AFFB025E CRC64;

  Query Match             55.4%;  Score 36;  DB 1;  Length 486;
  Best Local Similarity   58.3%;  Pred. No. 56;
  Matches    7;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 GYDDQESVKSKV 13
              ||| |:|  ||
Db        228 GYDHQDSTSLKV 239


RESULT 15
CHIB_SERMA
ID   CHIB_SERMA     STANDARD;      PRT;   499 AA.
AC   P11797;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Chitinase B precursor (EC 3.2.1.14).
GN   CHIB.
OS   Serratia marcescens.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Serratia.
OX   NCBI_TaxID=615;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 990 / QMB1466;
RX   MEDLINE=89345210; PubMed=2668886;
RA   Harpster M.H., Dunsmuir P.;
RT   "Nucleotide sequence of the chitinase B gene of Serratia marcescens
RT   QMB1466.";
RL   Nucleic Acids Res. 17:5395-5395(1989).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=2170;
RX   MEDLINE=98037511; PubMed=9371460;
RA   Watanabe T., Kimura K., Sumiya T., Nikaidou N., Suzuki K., Suzuki M.,
RA   Taiyoji M., Ferrer S., Reque M.;
RT   "Genetic analysis of the chitinase system of Serratia marcescens
RT   2170.";
RL   J. Bacteriol. 179:7111-7117(1997).
CC   -!- CATALYTIC ACTIVITY: Hydrolysis of the 1,4-beta-linkages of N-
CC       acetyl-D-glucosamine polymers of chitin.
CC   -!- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
CC       HYDROLASES).
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X15208; CAA33278.1; -.
DR   EMBL; AB015997; BAA31568.1; -.
DR   PIR; S04856; S04856.
DR   HSSP; P07254; 1CIN.
DR   InterPro; IPR003610; CBM_5_12.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   InterPro; IPR001579; Glyco_hydro_18/2.
DR   Pfam; PF02839; CBM_5_12; 1.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   ProDom; PD000471; Glyco_hydro_18; 1.
DR   SMART; SM00495; ChtBD3; 1.
DR   SMART; SM00636; Glyco_18; 1.
DR   PROSITE; PS01095; CHITINASE_18; 1.
KW   Hydrolase; Glycosidase; Chitin degradation; Signal.
FT   SIGNAL      1     41
FT   CHAIN      42    499       CHITINASE B.
FT   ACT_SITE   144    144       PROTON DONOR (BY SIMILARITY).
SQ   SEQUENCE   499 AA;  55464 MW;  FFD674916109D1D8 CRC64;

  Query Match             55.4%;  Score 36;  DB 1;  Length 499;
  Best Local Similarity   66.7%;  Pred. No. 57;
  Matches    8;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 VGYDDQESVKSK 12
              | |:| || | |
Db        377 VTYDDAESFKYK 388


Search completed: September 24, 2003, 13:42:29
Job time : 9.17978 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 24, 2003, 13:33:45 ; Search time 42.2135 Seconds
(without alignments)
79.469 Million cell updates/sec

Title: US-09-744-282-7
Perfect score: 65
Sequence: 1 VGYDDQESVKSKV 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	65	100.0	383	4	Q96HI7	Q96hi7 homo sapien
2	65	100.0	383	4	Q8IVA4	Q8iva4 homo sapien
3	58	89.2	383	6	Q29411	Q2941l sus scrofa
4	56	86.2	352	11	Q9WTV1	Q9wtv1 rattus norv
5	54	83.1	381	11	Q99J84	Q99j84 mus musculu
6	54	83.1	389	11	Q8BKL8	Q8bkl8 mus musculu
7	52	80.0	332	6	O18949	O18949 bos taurus
8	52	80.0	383	6	Q8SPQ0	Q8spq0 capra hircu
9	50	76.9	2838	5	Q8MP05	Q8mp05 tenebrio mo
10	48	73.8	371	10	Q81856	Q81856 arabidopsis
11	47	72.3	739	10	Q9SWX8	Q9swx8 nicotiana t
12	47	72.3	929	5	Q8MY79	Q8my79 haemaphysal
13	46	70.8	387	4	Q9H3V8	Q9h3v8 homo sapien
14	46	70.8	421	10	Q81853	Q81853 arabidopsis
15	46	70.8	466	4	Q13231	Q13231 homo sapien
16	45	69.2	379	10	Q81862	Q81862 arabidopsis
17	44	67.7	248	10	Q81858	Q81858 arabidopsis
18	44	67.7	332	10	Q81854	Q81854 arabidopsis
19	44	67.7	390	4	Q96F97	Q96f97 homo sapien
20	44	67.7	1080	5	Q9BLI6	Q9bli6 bombyx mori
21	43	66.2	366	10	Q81863	Q81863 arabidopsis
22	43	66.2	378	10	Q43591	Q43591 nicotiana t
23	43	66.2	525	5	O44079	O44079 anopheles g
24	43	66.2	983	5	Q9VZV2	Q9vzv2 drosophila
25	43	66.2	1013	5	Q960M0	Q960m0 drosophila
26	42	64.6	365	10	Q81857	Q81857 arabidopsis
27	42	64.6	460	5	Q8WS95	Q8ws95 glossina mo
28	42	64.6	488	13	Q90W34	Q90w34 bufo japoni
29	42	64.6	526	16	Q9PNK1	Q9pnk1 campylobact
30	42	64.6	633	5	Q22468	Q22468 caenorhabdi
31	42	64.6	1030	3	Q9HFQ9	Q9hfq9 emericella
32	41	63.1	319	5	Q97155	Q97155 toxoplasma
33	41	63.1	353	16	Q8RD13	Q8rd13 thermoanaer
34	41	63.1	440	2	Q8GI68	Q8gi68 clostridium
35	41	63.1	440	2	Q8GI67	Q8gi67 clostridium
36	41	63.1	440	2	Q8GI66	Q8gi66 clostridium
37	41	63.1	440	2	Q8GI65	Q8gi65 clostridium
38	41	63.1	440	2	Q8G8K0	Q8g8k0 clostridium
39	41	63.1	443	2	Q53482	Q53482 clostridium
40	41	63.1	477	5	Q966Z0	Q966z0 suberites d
41	41	63.1	1011	2	Q9L5M2	Q9l5m2 salmonella
42	41	63.1	1011	16	Q935H0	Q935h0 salmonella
43	40	61.5	131	5	Q8T5B8	Q8t5b8 plasmodium
44	40	61.5	287	2	Q8RSX8	Q8rsx8 bacillus th
45	40	61.5	349	17	Q980P6	Q980p6 sulfolobus

ALIGNMENTS

RESULT 1
Q96HI7
ID Q96HI7 PRELIMINARY; PRT; 383 AA.
AC Q96HI7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Similar to chitinase 3-like 1 (Cartilage glycoprotein-39).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC008568; AAH08568.1; -.
DR InterPro; IPR001223; Glyco_hydro_18.
DR InterPro; IPR005829; Sug_transporter.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR SMART; SM00636; Glyco_18; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
SQ SEQUENCE 383 AA; 42625 MW; 6C9BF133BDC7C2D1 CRC64;

Query Match 100.0%; Score 65; DB 4; Length 383;
Best Local Similarity 100.0%; Pred. No. 0.0034;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VGYDDQESVKSKV 13
|||||||||||||
Db 326 VGYDDQESVKSKV 338

RESULT 2
Q8IVA4
ID Q8IVA4 PRELIMINARY; PRT; 383 AA.
AC Q8IVA4;

```
DR   PROSITE; PS00165; DEHYDRATASE_SER_THR; 1.
KW   Signal; Hydrolase; Glycosidase.
FT   SIGNAL        1     20       POTENTIAL.
FT   CHAIN        21   2838       CHITINASE.
SQ   SEQUENCE   2838 AA;  321407 MW;  608B3F2A8E98B9B7 CRC64;

  Query Match             76.9%;  Score 50;  DB 5;  Length 2838;
  Best Local Similarity   76.9%;  Pred. No. 11;
  Matches   10;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 VGYDDQESVKSKV 13
              :|||| ||||: :|
Db        516 VGYDDAESVERKV 528


RESULT 10
O81856
ID   O81856      PRELIMINARY;      PRT;   371 AA.
AC   O81856;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Chitinase - like protein (Chitinase-like protein).
GN   CHIC1 OR T16H5.110 OR AT4G19750.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   De haan M., Maarse A.C., Grivell L.A., Bancroft I., Mewes H.W.,
RA   Mayer K., Schueller C., Bevan M.;
RL   Submitted (JUN-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   De Haan M., Maarse A.C., Grivell L.A., Mewes H.W., Lemcke K.,
RA   Mayer K.F.X.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   EU Arabidopsis sequencing project;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL024486; CAA19692.1; -.
DR   EMBL; AL161551; CAB78977.1; -.
DR   InterPro; IPR001064; Crystallin.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   ProDom; PD000471; Glyco_hydro_18; 1.
DR   SMART; SM00636; Glyco_18; 1.
DR   PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; 1.
SQ   SEQUENCE   371 AA;  40399 MW;  05871145067ACF42 CRC64;

  Query Match             73.8%;  Score 48;  DB 10;  Length 371;
  Best Local Similarity   61.5%;  Pred. No. 3.2;
  Matches    8;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 VGYDDQESVKSKV 13
              :||.| :|: :||
Db        311 IGYDDNQSIVSKV 323


RESULT 11
Q9SWX8
ID   Q9SWX8      PRELIMINARY;      PRT;   739 AA.
AC   Q9SWX8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Receptor-like kinase CHRK1.
GN   CHRK1.
OS   Nicotiana tabacum (Common tobacco).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; lamiids; Solanales; Solanaceae; Nicotiana.
OX   NCBI_TaxID=4097;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kim Y.-S., Yoon K.-M., Cho H.-S., Liu J.-R., Lee H.-S.;
RT   "CHRK1, a receptor-like kinase of tobacco has a chitinase domain in
RT   its N terminus.";
RL   Submitted (SEP-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR   EMBL; AF088885; AAD52097.1; -.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   Pfam; PF00069; pkinase; 1.
DR   ProDom; PD000471; Glyco_hydro_18; 2.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00636; Glyco_18; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ   SEQUENCE   739 AA;  83977 MW;  4763B79486D346E3 CRC64;

  Query Match             72.3%;  Score 47;  DB 10;  Length 739;
  Best Local Similarity   53.8%;  Pred. No. 9.5;
  Matches    7;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 VGYDDQESVKSKV 13
              :|.|| |::::||
Db        306 IGYDDVEAIRTKV 318


RESULT 12
Q8MY79
ID   Q8MY79      PRELIMINARY;      PRT;   929 AA.
AC   Q8MY79;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Chitinase.
GN   CHT.
OS   Haemaphysalis longicornis.
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC   Parasitiformes; Ixodida; Ixodidae; Haemaphysalis.
OX   NCBI_TaxID=44386;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   You M.;
RT   "Molecular characterization of a chitinase protein from the hard tick
RT   Haemaphysalis longicornis.";
RL   Submitted (NOV-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB074977; BAC06447.1; -.
DR   InterPro; IPR002557; Chitin_bind_PerA.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   InterPro; IPR001579; Glyco_hydro_18/2.
DR   Pfam; PF01607; CBM_14; 1.
DR   Pfam; PF00704; Glyco_hydro_18; 2.
DR   ProDom; PD000471; Glyco_hydro_18; 2.
DR   SMART; SM00494; ChtBD2; 1.
DR   SMART; SM00636; Glyco_18; 2.
DR   PROSITE; PS01095; CHITINASE_18; 2.
SQ   SEQUENCE   929 AA;  104423 MW;  3D70C955DC1D93C6 CRC64;

  Query Match             72.3%;  Score 47;  DB 5;  Length 929;
  Best Local Similarity   69.2%;  Pred. No. 12;
  Matches    9;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 VGYDDQESVKSKV 13
              :|:|:| |:| :|
Db        786 VGFDDQRSLKLKV 798
```

```
RESULT 13
Q9H3V8
ID   Q9H3V8      PRELIMINARY;      PRT:   387 AA.
AC   Q9H3V8;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Chitotriosidase precursor.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96064695; PubMed=7592832;
RA   Boot R.G., Renkema G.H., Strijland A., van Zonneveld A.J., Aerts J.M.;
RT   "Cloning of a cDNA encoding chitotriosidase, a human chitinase
RT   produced by macrophages.";
RL   J. Biol. Chem. 270:26252-26256(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98421482; PubMed=9748235;
RA   Boot R.G., Renkema G.H., Verhoek M., Strijland A., Bliek J.,
RA   de Meulemeester T.M., Mannens M.M., Aerts J.M.;
RT   "The human chitotriosidase gene. Nature of inherited enzyme
RT   deficiency.";
RL   J. Biol. Chem. 273:25680-25685(1998).
DR   EMBL; U62662; AAG10644.1; -.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   InterPro; IPR001579; Glyco_hydro_18/2.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   ProDom; PD000471; Glyco_hydro_18; 1.
DR   SMART; SM00636; Glyco_18; 1.
DR   PROSITE; PS01095; CHITINASE_18; 1.
KW   Signal.
FT   SIGNAL        1     21       POTENTIAL.
FT   CHAIN        22    387       CHITOTRIOSIDASE.
SQ   SEQUENCE   387 AA;  43133 MW;  03A272B8BC5E0D71 CRC64;

  Query Match             70.8%;  Score 46;  DB 4;  Length 387;
  Best Local Similarity   69.2%;  Pred. No. 7.4;
  Matches    9;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 VGYDDQESVKSKV 13
              ||:|| || :||
Db        332 VGFDDVESFKTKV 344


RESULT 14
O81853
ID   O81853      PRELIMINARY;      PRT:   421 AA.
AC   O81853;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical 46.9 kDa protein.
GN   T16H5.80 OR AT4G19720.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   De haan M., Maarse A.C., Grivell L.A., Bancroft I., Mewes H.W.,
RA   Mayer K., Schueller C., Bevan M.;
RL   Submitted (JUN-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   De Haan M., Maarse A.C., Grivell L.A., Mewes H.W., Lemcke K.;
RA   Mayer K.F.X.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   EU Arabidopsis sequencing project;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL024486; CAA19689.1; -.
DR   EMBL; AL161551; CAB78974.1; -.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   InterPro; IPR000531; TonB_boxC.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   ProDom; PD000471; Glyco_hydro_18; 2.
DR   SMART; SM00636; Glyco_18; 1.
DR   PROSITE; PS00430; TONB_DEPENDENT_REC_1; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   421 AA;  46895 MW;  4705B6C1C68DB7B9 CRC64;

  Query Match             70.8%;  Score 46;  DB 10;  Length 421;
  Best Local Similarity   61.5%;  Pred. No. 8.1;
  Matches    8;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 VGYDDQESVKSKV 13
              :|||: :|| :||
Db        364 IGYDDHQSVVAKV 376


RESULT 15
Q13231
ID   Q13231      PRELIMINARY;      PRT:   466 AA.
AC   Q13231;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Chitotriosidase precursor.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96064695; PubMed=7592832;
RA   Boot R.G., Renkema G.H., Strijland A., van Zonneveld A.J., Aerts J.M.;
RT   "Cloning of a cDNA encoding chitotriosidase, a human chitinase
RT   produced by macrophages.";
RL   J. Biol. Chem. 270:26252-26256(1995).
DR   EMBL; U29615; AAC50246.1; -.
DR   Genew; HGNC:1936; CHIT1.
DR   InterPro; IPR002557; Chitin_bind_PerA.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   InterPro; IPR001579; Glyco_hydro_18/2.
DR   Pfam; PF01607; CBM_14; 1.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   ProDom; PD000471; Glyco_hydro_18; 1.
DR   SMART; SM00494; ChtBD2; 1.
DR   SMART; SM00636; Glyco_18; 1.
DR   PROSITE; PS01095; CHITINASE_18; 1.
KW   Signal.
FT   SIGNAL        1     21       POTENTIAL.
FT   CHAIN        22    466       CHITOTRIOSIDASE.
SQ   SEQUENCE   466 AA;  51681 MW;  B4312D1E885E386D CRC64;

  Query Match             70.8%;  Score 46;  DB 4;  Length 466;
  Best Local Similarity   69.2%;  Pred. No. 8.9;
  Matches    9;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 VGYDDQESVKSKV 13
              ||:|| || |:||
Db        332 VGFDDVESFKTKV 344


Search completed: September 24, 2003, 13:47:27
Job time : 43.2135 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 24, 2003, 13:25:37 ; Search time 56.1573 Seconds
(without alignments)
39.570 Million cell updates/sec

Title: US-09-744-282-8
Perfect score: 66
Sequence: 1 SQRFSKIASNTQSR 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 156726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	66	100.0	14	17	AAR95078	Peptide contg. hC
2	66	100.0	14	21	AAY69233	Peptide fragment o
3	66	100.0	14	23	AAE20223	Human cartilage gl
4	66	100.0	105	22	AAU02074	Synthetic human ta
5	66	100.0	126	21	AAG00227	Human secreted pro
6	66	100.0	220	22	AAU02078	Synthetic human mu
7	66	100.0	281	22	AAU02079	Synthetic multi-ta
8	66	100.0	308	24	ABR41620	Human DITHP secret
9	66	100.0	321	9	AAP81342	Polypeptide involv
10	66	100.0	350	22	AAU02076	Synthetic human mu
11	66	100.0	383	21	AAB03442	Gp38k protein sequ
12	66	100.0	383	23	ABG36297	Human ovarian canc
13	66	100.0	383	24	ABU56651	Lung cancer-associ
14	66	100.0	459	22	AAU02077	Synthetic multi-ta
15	61	92.4	16	18	AAW45113	Human cartilage gl
16	56	84.8	383	18	AAW26751	Bovine whey protei
17	52	78.8	16	18	AAW45114	Human cartilage gl
18	43	65.2	9	17	AAR95074	Human cartilage gl
19	43	65.2	9	21	AAY69229	Peptide fragment o
20	40	60.6	794	24	ABJ25507	Aspergillus fumiga
21	40	60.6	1138	24	ABJ26107	Aspergillus fumiga
22	37	56.1	522	21	AAG29819	Arabidopsis thalia
23	37	56.1	526	21	AAG29818	Arabidopsis thalia
24	37	56.1	577	21	AAG29817	Arabidopsis thalia
25	37	56.1	1216	22	ABB60349	Drosophila melanog
26	36	54.5	78	22	AAG74411	Human colon cancer
27	36	54.5	150	22	ABG07997	Novel human diagno
28	36	54.5	525	22	ABG16066	Novel human diagno
29	36	54.5	1817	22	ABB71955	Drosophila melanog
30	35	53.0	138	23	ABU50696	Helicobacter pylor
31	35	53.0	150	22	AAU55705	Propionibacterium
32	35	53.0	267	21	AAG08559	Arabidopsis thalia
33	35	53.0	269	21	AAG08558	Arabidopsis thalia
34	35	53.0	285	21	AAG08557	Arabidopsis thalia
35	35	53.0	371	23	A5B07726	Sunflower chitinas
36	35	53.0	485	21	AAG30001	Arabidopsis thalia
37	35	53.0	485	21	AAG51798	Arabidopsis thalia
38	35	53.0	495	23	ABR83234	Herbicidally activ
39	35	53.0	524	21	AAG30000	Arabidopsis thalia
40	35	53.0	524	21	AAG51797	Arabidopsis thalia
41	35	53.0	528	21	AAG29999	Arabidopsis thalia
42	35	53.0	531	21	AAG51796	Arabidopsis thalia
43	35	53.0	1050	22	ABG24938	Novel human diagno
44	35	53.0	1230	19	AAW98275	H. pylori GHPO 690
45	35	53.0	1237	18	AAW55640	H. pylori ORF 04ep

ALIGNMENTS

RESULT 1
AAR95078
ID AAR95078 standard; peptide; 14 AA.
XX
AC AAR95078;
XX
DT 17-DEC-1996 (first entry)
XX
DE Peptide contg. hC gp-39 autoantigen derived peptide (IV).
XX
KW Human; cartilage; glycoprotein 39; hC gp-39; autoantigen;
KW induction; T-cell; tolerance; mediated; destruction;
KW rheumatoid arthritis; disease model; diagnosis.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 4..12
FT /note= "HC gp-39 peptide (IV)"
XX
PN WO9613517-A1.
XX
PD 09-MAY-1996.
XX
PF 25-OCT-1995; 95WO-EP04201.
XX
PR 07-APR-1995; 95EP-0200886.
PR 27-OCT-1994; 94EP-0203128.
XX
PA (ALKU) AKZO NOBEL NV.
XX
PI Boots AMH, Verheijden GFM;

XX
DR WPI; 1996-239447/24.
XX
PT Peptide(s) derived from human cartilage gp-39 autoantigen - used
PT to induce specific T cell tolerance, esp. for treatment of
PT rheumatoid arthritis, also useful in producing arthritic animal
PT models
XX
PS Claim 2; Page 28; 37pp; English.
XX
CC The present peptide, contg. a peptide derived from the human
CC cartilage glycoprotein 39 (hC gp-39) autoantigen, can be used to
CC induce specific T-cell tolerance to hC gp-39 in a patient with
CC T-cell mediated cartilage destruction, esp. rheumatoid arthritis
CC (RA). It may also be used to induce RA in mice, etc. (human disease
CC models), and to diagnose the presence of activated T-cells
CC autoreactive with hC gp-39, i.e. to identify patients suitable for
CC tolerising treatment.
CC Therapeutically the peptide is given in a dose of 0.1-100 microg/kg
CC by injection, or 10-50 microg/kg to produce a disease model. The
CC peptide has a specific tolerising effect on autoreactive T-cells,
CC but leaves other components of the immune system intact, and
CC therefore should have few side effects.
XX
SQ Sequence 14 AA;

Query Match 100.0%; Score 66; DB 17; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.1e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SQRFSKIASNTQSR 14
||||||||||||||
Db 1 SQRFSKIASNTQSR 14

RESULT 2
AAY69233
ID AAY69233 standard; peptide; 14 AA.
XX
AC AAY69233;
XX
DT 30-MAY-2000 (first entry)
XX
DE Peptide fragment of human cartilage gp-39 (HC gp-39) protein.
XX
KW Human cartilage gp-39 protein; HC gp-39 protein; T-cell activator;
KW inflammatory disease; lymphocyte; antigen; bystander suppression;
KW antigen-specific T-cell tolerance; autoimmune disorder;
KW inflammatory disease; rheumatoid arthritis;
KW autoimmune inflammatory disease; autoantigen.
XX
OS Homo sapiens.
XX
PN WO200004917-A2.
XX
PD 03-FEB-2000.
XX
PF 19-JUL-1999; 99WO-EP05331.
XX
PR 23-JUL-1998; 98EP-0202471.
XX
PA (ALKU) AKZO NOBEL NV.
XX
PI Miltenburg AMM, Boots AMH;
XX
DR WPI; 2000-182537/16.
XX
PT Use of cartilage protein or peptides for treating inflammatory diseases
PT especially rheumatoid arthritis -
XX
PS Claim 2; Page 17; 29pp; English.
XX
CC AAY69226-33 represent fragments of the human cartilage gp-39 (HC gp-39)
CC protein. The peptides are T-cell activators. The peptides are used for
CC preventing inflammatory diseases, by modulating lymphocytes reactive
CC to antigens other than HC gp-39, present in the same tissue of
CC HC gp-39. By the induction of antigen-specific T-cell tolerance,
CC autoimmune disorders can be treated by bystander suppression. The
CC HC gp-39 peptides are useful for preventing inflammatory diseases
CC especially rheumatoid arthritis. The fragments may also be used for
CC treating autoimmune inflammatory diseases. The HC gp-39 peptides are
CC especially beneficial in autoimmune conditions in which the
CC autoantigens associated with the condition are unknown.
XX
SQ Sequence 14 AA;

Query Match 100.0%; Score 66; DB 21; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.1e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SQRFSKIASNTQSR 14
||||||||||||||
Db 1 SQRFSKIASNTQSR 14

RESULT 3
AAE20223
ID AAE20223 standard; peptide; 14 AA.
XX
AC AAE20223;
XX
DT 18-JUN-2002 (first entry)
XX
DE Human cartilage glycoprotein (HC gp)-39 peptide (residues 103-116).
XX
KW Human cartilage glycoprotein-39; HC gp-39; autoimmune disorder; MHC;
KW major histocompatibility complex; insulin-dependent diabetes mellitus;
KW multiple sclerosis; psoriasis; myasthenia gravis; rheumatoid arthritis;
KW immunosuppressive; neuroprotective; therapy.
XX
OS Homo sapiens.
XX
PN WO200214870-A2.
XX
PD 21-FEB-2002.
XX
PF 08-AUG-2001; 2001WO-EP09136.
XX
PR 14-AUG-2000; 2000EP-0202844.
XX
PA (ALKU) AKZO NOBEL NV.
XX
PI Steenbakkers PGA;
XX
DR WPI; 2002-269214/31.
XX
PT Method for diagnosing an autoimmune disease activity by detecting the
PT presence of an autoimmune specific major histocompatibility complex
PT (MHC)-peptide complex in a patient using antibodies specific against
PT MHC-peptide complexes -
XX
PS Disclosure; Page 43; 43pp; English.
XX
CC The patent discloses a method for diagnosing an autoimmune disease
CC activity. The method involves detecting the presence of an autoimmune
CC specific major histocompatibility complex (MHC)-peptide complex in a
CC patient suffering from an autoimmune disease with antibodies or its
CC antigen binding domains that specifically bind to a MHC-gp39-derived
CC peptide. The method and antibodies are useful for diagnosing auto-
CC immune diseases or purifying autoimmune specific MHC-peptide complexes.
CC They are also useful in therapy or for the manufacture of pharmaceutical
CC preparation for the treatment of autoimmune disorders such as insulin-
CC dependent diabetes mellitus, multiple sclerosis, myasthenia gravis,
CC psoriasis or rheumatoid arthritis. The present sequence is human
CC cartilage glycoprotein (HC gp)-39 peptide (residues 103-116). This
CC peptide is used in the exemplification of the invention.

```
XX
SQ   Sequence   14 AA;

  Query Match             100.0%;  Score 66;  DB 23;  Length 14;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-05;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SQRFSKIASNTQSR 14
              ||||||||||||||
Db          1 SQRFSKIASNTQSR 14


RESULT 4
AAU02074
ID   AAU02074 standard; Protein; 105 AA.
XX
AC   AAU02074;
XX
DT   07-SEP-2001  (first entry)
XX
DE   Synthetic human target autoantigen shGLP/RA.
XX
KW   shGLP/RA; immunogenic epitope cluster; IEC;
KW   synthetic human chondrocyte glycoprotein 69/cartilage link protein;
KW   autoantigen; autoimmune disease; multiple sclerosis;
KW   rheumatoid arthritis; insulin-dependent diabetes mellitus;
KW   myasthenia gravis; uveitis; autoimmune hepatitis; thyroiditis; insulitis;
KW   orchitis; idiopathic thrombocytopenic purpura; inflammatory disease;
KW   Crohn's disease; ulcerative colitis.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 103
FT                   /label= OTHER
FT                   /note= "Other: STOP"
XX
PN   WO200131037-A2.
XX
PD   03-MAY-2001.
XX
PF   26-OCT-2000; 2000WO-IL00688.
XX
PR   27-OCT-1999;   99IL-0132611.
XX
PA   (YEDA ) YEDA RES & DEV CO LTD.
XX
PI   Ben-Nun A,  Kerlero De Rosbo N,  Sappler GP;
XX
DR   WPI; 2001-300515/31.
DR   N-PSDB; AAS04827.
XX
PT   Novel synthetic human target autoantigen gene useful for treating
PT   autoimmune diseases such as multiple sclerosis, insulin-dependent
PT   diabetes mellitus, rheumatoid arthritis, myasthenia gravis, and uveitis
PT   -
XX
PS   Claim 173; Fig 51; 182pp; English.
XX
CC   The sequence represents shGLP/RA (synthetic human chondrocyte
CC   glycoprotein 69/cartilage link protein) containing immunogenic epitope
CC   clusters (IEC) from chondrocyte glycoprotein 69 and cartilage link
CC   protein. The synthetic human target autoantigen genes of the invention
CC   comprise sequences coding for at least 2 IECs of autoantigen(s) related
CC   to a specific autoimmune disease. The synthetic human target autoantigen
CC   genes are useful for treating autoimmune diseases such as multiple
CC   sclerosis, insulin-dependent diabetes mellitus, rheumatoid arthritis,
CC   myasthenia gravis, uveitis, autoimmune hepatitis, thyroiditis, insulitis,
CC   orchitis, idiopathic thrombocytopenic purpura, and inflammatory diseases
CC   (Crohn's disease, ulcerative colitis). The synthetic human target
CC   autoantigen genes are also useful for diagnosis and/or monitoring the
CC   progression of the autoimmune disease.
XX
SQ   Sequence   105 AA;

  Query Match             100.0%;  Score 66;  DB 22;  Length 105;
  Best Local Similarity   100.0%;  Pred. No. 0.0001;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SQRFSKIASNTQSR 14
              ||||||||||||||
Db          8 SQRFSKIASNTQSR 21


RESULT 5
AAG00227
ID   AAG00227 standard; Protein; 126 AA.
XX
AC   AAG00227;
XX
DT   06-OCT-2000  (first entry)
XX
DE   Human secreted protein, SEQ ID NO: 4308.
XX
KW   Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW   gene therapy; chromosome mapping.
XX
OS   Homo sapiens.
XX
PN   EP1033401-A2.
XX
PD   06-SEP-2000.
XX
PF   21-FEB-2000; 2000EP-0200610.
XX
PR   26-FEB-1999;   99US-0122487.
XX
PA   (GEST ) GENSET.
XX
PI   Dumas Milne Edwards J,  Duclert A,  Giordano J;
XX
DR   WPI; 2000-500381/45.
DR   N-PSDB; AAC00233.
XX
PT   New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT   obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT   diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
PS   Claim 13; SEQ ID 4308; 71pp + CD-ROM; English.
XX
CC   The present sequence is a polypeptide encoded by one of a large number
CC   of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
CC   were prepared from total human RNAs or polyA+ RNAs derived from 30
CC   different tissues. EST sequences usually correspond mainly to the 3'
CC   untranslated region (UTR) of the mRNA because they are often obtained
CC   from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
CC   isolating cDNA sequences derived from the 5' ends of mRNAs and even in
CC   those cases where longer cDNA sequences have been obtained, the full 5'
CC   UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
CC   ends and can therefore be used to obtain full length cDNAs and genomic
CC   DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
CC   chromosome mapping procedures. They are used to obtain upstream
CC   regulatory sequences and to design expression and secretion vectors.
XX
SQ   Sequence   126 AA;

  Query Match             100.0%;  Score 66;  DB 21;  Length 126;
  Best Local Similarity   100.0%;  Pred. No. 0.00013;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SQRFSKIASNTQSR 14
              ||||||||||||||
Db        103 SQRFSKIASNTQSR 116
```

```
RESULT 6
AAU02078
ID   AAU02078 standard; Protein; 220 AA.
XX
AC   AAU02078;
XX
DT   07-SEP-2001  (first entry)
XX
DE   Synthetic human multi-target autoantigen Y-RAPc.
XX
KW   Y-RAPc; immunogenic epitope cluster; IEC;
KW   autoantigen; shCollagen/RA; shAggrecan/RA; shGLP/RA;
KW   autoimmune disease; multiple sclerosis; rheumatoid arthritis;
KW   insulin-dependent diabetes mellitus; myasthenia gravis; uveitis;
KW   autoimmune hepatitis; thyroiditis; insulitis; orchitis;
KW   idiopathic thrombocytopenic purpura; inflammatory disease;
KW   Crohn's disease; ulcerative colitis; shMultiTAG.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 218
FT                   /label= OTHER
FT                   /note= "Other= STOP"
XX
PN   WO200131037-A2.
XX
PD   03-MAY-2001.
XX
PF   26-OCT-2000; 2000WO-IL00688.
XX
PR   27-OCT-1999;   99IL-0132611.
XX
PA   (YEDA ) YEDA RES & DEV CO LTD.
XX
PI   Ben-Nun A,  Kerlero De Rosbo N,  Sappler GP;
XX
DR   WPI; 2001-300515/31.
DR   N-PSDB; AAS04839.
XX
PT   Novel synthetic human target autoantigen gene useful for treating
PT   autoimmune diseases such as multiple sclerosis, insulin-dependent
PT   diabetes mellitus, rheumatoid arthritis, myasthenia gravis, and uveitis
PT   .
XX
PS   Claim 185; Fig 59; 182pp; English.
XX
CC   The sequence represents synthetic human multi-target autoantigen
CC   (shMultiTAG) Y-RAPa consisting of regions encoding preferred epitopes
CC   from shCollagen/RA, shAggrecan/RA and shGLP/RA. The synthetic human
CC   target autoantigen genes of the invention comprise sequences coding for
CC   at least 2 IECs of autoantigen(s) related to a specific autoimmune
CC   disease. The synthetic human target autoantigen genes are useful for
CC   treating autoimmune diseases such as multiple sclerosis, insulin-
CC   dependent diabetes mellitus, rheumatoid arthritis, myasthenia gravis,
CC   uveitis, autoimmune hepatitis, thyroiditis, insulitis, orchitis,
CC   idiopathic thrombocytopenic purpura, and inflammatory diseases (Crohn's
CC   disease, ulcerative colitis). The synthetic human target autoantigen
CC   genes are also useful for diagnosis and/or monitoring the progression of
CC   the autoimmune disease.
XX
SQ   Sequence   220 AA;

  Query Match           100.0%;  Score 66;  DB 22;  Length 220;
  Best Local Similarity  100.0%;  Pred. No. 0.00024;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SQRFSKIASNTQSR 14
              ||||||||||||||
Db        164 SQRFSKIASNTQSR 177
```

```
RESULT 7
AAU02079
ID   AAU02079 standard; Protein; 281 AA.
XX
AC   AAU02079;
XX
DT   07-SEP-2001  (first entry)
XX
DE   Synthetic multi-target autoantigen Y-RAPd.
XX
KW   Y-RAPd; immunogenic epitope cluster; IEC;
KW   autoantigen; shCollagen/RA; shAggrecan/RA; shGLP/RA; shHSP/RA;
KW   autoimmune disease; multiple sclerosis; rheumatoid arthritis;
KW   insulin-dependent diabetes mellitus; myasthenia gravis; uveitis;
KW   autoimmune hepatitis; thyroiditis; insulitis; orchitis;
KW   idiopathic thrombocytopenic purpura; inflammatory disease;
KW   Crohn's disease; ulcerative colitis; shMultiTAG.
XX
OS   Synthetic.
OS   Homo sapiens.
OS   Escherichia coli.
OS   Mycobacterium tuberculosis.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 279
FT                   /label= OTHER
FT                   /note= "Other= STOP"
XX
PN   WO200131037-A2.
XX
PD   03-MAY-2001.
XX
PF   26-OCT-2000; 2000WO-IL00688.
XX
PR   27-OCT-1999;   99IL-0132611.
XX
PA   (YEDA ) YEDA RES & DEV CO LTD.
XX
PI   Ben-Nun A,  Kerlero De Rosbo N,  Sappler GP;
XX
DR   WPI; 2001-300515/31.
DR   N-PSDB; AAS04840.
XX
PT   Novel synthetic human target autoantigen gene useful for treating
PT   autoimmune diseases such as multiple sclerosis, insulin-dependent
PT   diabetes mellitus, rheumatoid arthritis, myasthenia gravis, and uveitis
PT   .
XX
PS   Claim 188; Fig 61; 182pp; English.
XX
CC   The sequence represents synthetic human multi-target autoantigen
CC   (shMultiTAG) Y-RAPd consisting of preferred epitopes encoded by
CC   shCollagen/RA, shAggrecan/RA, shGLP/RA and shHSP/RA. The synthetic human
CC   target autoantigen genes of the invention comprise sequences coding for
CC   at least 2 IECs of autoantigen(s) related to a specific autoimmune
CC   disease. The synthetic human target autoantigen genes are useful for
CC   treating autoimmune diseases such as multiple sclerosis, insulin-
CC   dependent diabetes mellitus, rheumatoid arthritis, myasthenia gravis,
CC   uveitis, autoimmune hepatitis, thyroiditis, insulitis, orchitis,
CC   idiopathic thrombocytopenic purpura, and inflammatory diseases (Crohn's
CC   disease, ulcerative colitis). The synthetic human target autoantigen
CC   genes are also useful for diagnosis and/or monitoring the progression of
CC   the autoimmune disease.
XX
SQ   Sequence   281 AA;

  Query Match           100.0%;  Score 66;  DB 22;  Length 281;
  Best Local Similarity  100.0%;  Pred. No. 0.00031;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SQRFSKIASNTQSR 14
              ||||||||||||||
Db        164 SQRFSKIASNTQSR 177
```

```
RESULT 8
ABR41620
ID   ABR41620 standard; Protein; 308 AA.
XX
AC   ABR41620;
XX
DT   02-JUN-2003  (first entry)
XX
DE   Human DITHP secreted/extracellular matrix protein.
XX
KW   Human; dithp; diagnostic and therapeutic polynucleotide; diagnosis;
KW   cancer; cell proliferative disorder; autoimmune disorder;
KW   inflammatory disorder; infection; hormonal disorder; metabolic disorder;
KW   neurological disorder; gastrointestinal disorder; transport disorder;
KW   connective tissue disorder; drug screening; proteome analysis;
KW   gene therapy; antisense therapy; genotyping; transgenic animal; knock in;
KW   disease model; toxicological testing; transcript imaging;
KW   secreted protein; extracellular matrix.
XX
OS   Homo sapiens.
XX
PN   WO200297031-A2.
XX
PD   05-DEC-2002.
XX
PF   27-MAR-2002; 2002WO-US10056.
XX
PR   28-MAR-2001; 2001US-279619P.
PR   29-MAR-2001; 2001US-280067P.
PR   29-MAR-2001; 2001US-280068P.
PR   16-MAY-2001; 2001US-291280P.
PR   17-MAY-2001; 2001US-291829P.
PR   17-MAY-2001; 2001US-291849P.
PR   19-JUN-2001; 2001US-299428P.
PR   20-JUN-2001; 2001US-299776P.
PR   20-JUN-2001; 2001US-300001P.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Datta A,  Jones AE,  Tran AB,  Dahl CE,  Gietzen D,  Chinn J;
PI   Dufour GE,  Hillman JL,  Yue JY,  Tuason O,  Yao FE,  Ansbey SR;
PI   Daughtery SC,  Dam IC,  Liu TF,  Nguyen DA,  Kloefeld Y,  Gerstin EH;
PI   Peralta CH,  David MH,  Lewis SA,  Chen AJ,  Panzer SR,  Harris B;
PI   Flores V,  Marwaha R,  Lo A,  Lam RY,  Urashka MB;
XX
DR   WPI; 2003-129518/12.
DR   N-PSDB; ACC46557.
XX
PT   Novel human diagnostic and therapeutic polypeptide useful for
PT   identifying test compound which specifically binds to a polypeptide
PT   encoded by human diagnostic and therapeutic polynucleotide, and to
PT   induce antibodies  -
XX
PS   Claim 27; SEQ ID No 1155; 591pp; English.
XX
CC   The invention relates to novel human diagnostic and therapeutic
CC   polynucleotides designated dithp (ACC46080-ACC46749) and to their
CC   encoded proteins (DITHP; ABR41136-ABR41812). The invention also relates
CC   to polynucleotide sequences at least 90% identical to the dithp cDNA
CC   sequences of the invention; recombinant vectors, host cells and
CC   transgenic organisms comprising a dithp nucleic acid sequence; the
CC   recombinant production of DITHP proteins; antibodies specific for DITHP
CC   proteins; microarrays comprising dithp nucleic acid sequences; methods
CC   of detecting dithp nuclectide and protein sequences; methods of screening
CC   for compounds which specifically bind a DITHP protein; and methods of
CC   assessing the toxicity of test compounds using a dithp hybridisation
CC   probe. Dithp nucleic acid sequences and DITHP proteins may be used in the
CC   diagnosis of a wide variety of conditions including cancer and other cell
CC   proliferative disorders; autoimmune or inflammatory disorders; bacterial,
CC   viral, fungal or parasitic infections; hormonal disorders; metabolic
CC   disorders; neurological disorders; gastrointestinal disorders; transport
CC   disorders; and connective tissue disorders. They may also be used to
CC   screen for modulators of protein activity or gene expression. DITHP
CC   proteins can additionally be used in analysis of the proteome of a tissue
CC   or cell type and to induce antibodies. The dithp nucleic acids are
CC   additionally useful in somatic or germline gene therapy of the disorders
CC   mentioned above, as a source of antisense sequences, as a source of
CC   probes and primers, in genotyping and identification of individuals, in
CC   the generation of transgenic animal models of human disease or knock in
CC   humanised animals, in toxicological testing, and in transcript imaging.
CC   The present sequence represents a DITHP protein which is a secreted/
CC   extracellular matrix protein.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   308 AA;

  Query Match             100.0%;  Score 66;  DB 24;  Length 308;
  Best Local Similarity   100.0%;  Pred. No. 0.00039;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SQRFSKIASNTQSR 14
              ||||||:||||||
Db        148 SQRFSKIASNTQSR 161


RESULT 9
AAP81342
ID   AAP81342 standard; protein; 321 AA.
XX
AC   AAP81342;
XX
DT   25-MAR-2003  (updated)
DT   10-MAR-2003  (updated)
DT   19-OCT-1990  (first entry)
XX
DE   Polypeptide involved in protective mechanisms.
XX
KW   Immune response; cell growth.
XX
OS   Unidentified.
XX
PN   JP63023898-A.
XX
PD   30-JAN-1988.
XX
PF   16-JUL-1986;   86JP-0167518.
XX
PR   16-JUL-1986;   86JP-0167518.
XX
PA   (DAIN ) DAINIPPON PHARM KK.
XX
DR   WPI; 1988-068419/10.
DR   P-PSDB; AAP81342.
XX
PT   New polypeptide and DNA encoding it - related to protective
PT   mechanisms such as immune response etc.
XX
PS   Disclosure; Page ?; ?pp; Japanese.
XX
CC   This polypeptide is involved in protective mechanisms such as immune
CC   response, cell growth and activation of protective functions.
CC   (Updated on 10-MAR-2003 to add missing OS field.)
CC   (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence   321 AA;

  Query Match             100.0%;  Score 66;  DB 9;  Length 321;
  Best Local Similarity   100.0%;  Pred. No. 0.00037;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SQRFSKIASNTQSR 14
              ||||||||||||||
```

```
Db        102 SQRFSKIASNTQSR 115


RESULT 10
AAU02076
ID   AAU02076 standard; Protein; 350 AA.
XX
AC   AAU02076;
XX
DT   07-SEP-2001  (first entry)
XX
DE   Synthetic human multi-target autoantigen Y-RAPa.
XX
KW   Y-RAPa; immunogenic epitope cluster; IEC;
KW   autoantigen; shCollagen/RA; shAggrecan/RA; shGLP/RA;
KW   autoimmune disease; multiple sclerosis; rheumatoid arthritis;
KW   insulin-dependent diabetes mellitus; myasthenia gravis; uveitis;
KW   autoimmune hepatitis; thyroiditis; insulitis; orchitis;
KW   idiopathic thrombocytopenic purpura; inflammatory disease;
KW   Crohn's disease; ulcerative colitis; shMultiTAG.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 348
FT                   /label= OTHER
FT                   /note= "Other- STOP"
XX
PN   WO200131037-A2.
XX
PD   03-MAY-2001.
XX
PF   26-OCT-2000; 2000WO-IL00688.
XX
PR   27-OCT-1999;   99IL-0132611.
XX
PA   (YEDA ) YEDA RES & DEV CO LTD.
XX
PI   Ben-Nun A,  Kerlero De Rosbo N,  Sappler GP;
XX
DR   WPI; 2001-300515/31.
DR   N-PSDB; AAS04857.
XX
PT   Novel synthetic human target autoantigen gene useful for treating
PT   autoimmune diseases such as multiple sclerosis, insulin-dependent
PT   diabetes mellitus, rheumatoid arthritis, myasthenia gravis, and uveitis
PT   -
XX
PS   Claim 179; Fig 55; 182pp; English.
XX
CC   The sequence represents synthetic human multi-target autoantigen
CC   (shMultiTAG) Y-RAPa consisting of regions encoding shCollagen/RA,
CC   shAggrecan/RA and shGLP/RA. The synthetic human target autoantigen genes
CC   of the invention comprise sequences coding for at least 2 IECs of
CC   autoantigen(s) related to a specific autoimmune disease. The synthetic
CC   human target autoantigen genes are useful for treating autoimmune
CC   diseases such as multiple sclerosis, insulin-dependent diabetes
CC   mellitus, rheumatoid arthritis, myasthenia gravis, uveitis, autoimmune
CC   hepatitis, thyroiditis, insulitis, orchitis, idiopathic thrombocytopenic
CC   purpura, and inflammatory diseases (Crohn's disease, ulcerative colitis).
CC   The synthetic human target autoantigen genes are also useful for
CC   diagnosis and/or monitoring the progression of the autoimmune disease.
XX
SQ   Sequence   350 AA;

  Query Match             100.0%;  Score 66;  DB 22;  Length 350;
  Best Local Similarity   100.0%;  Pred. No. 0.0004;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SQRFSKIASNTQSR 14
              ||||||||||||||
Db        253 SQRFSKIASNTQSR 266


RESULT 11
AAB03442
ID   AAB03442 standard; Protein; 383 AA.
XX
AC   AAB03442;
XX
DT   03-JAN-2001  (first entry)
XX
DE   Gp38k protein sequence.
XX
KW   Gp38k; chemoattractant; cell migration; wound healing; angiogenesis;
KW   cancer; vascular trauma; vascular disease; atherosclerosis; restenosis;
XX
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..24
FT                   /label= signal_peptide
FT   Region          30..37
FT                   /note= "peptide antibody"
FT   Modified-site   59..61
FT                   /label= glycosylation_site
FT   Binding-site    68..75
FT                   /label= leucine_zipper
FT   Binding-site    82..89
FT                   /label= leucine_zipper
FT   Active-site     131..136
FT                   /label= chitanase_active_site
FT   Binding-site    143..146
FT                   /label= heparin_binding_site
FT   Binding-site    147..154
FT                   /label= hyaluronic_acid_binding_site
FT   Binding-site    262..270
FT                   /label= hyaluronic_acid_binding_site
FT   Binding-site    278..281
FT                   /label= glycosaminoglycan_binding_site
FT   Region          354..357
FT                   /note= "acidic region"
FT   Binding-site    368..376
FT                   /label= hyaluronic_acid_binding_site
XX
PN   WO200034469-A1.
XX
PD   15-JUN-2000.
XX
PF   10-DEC-1999;   99WO-US29262.
XX
PR   11-DEC-1998;   98US-0111856.
XX
PA   (UYNY ) UNIV NEW YORK STATE RES FOUND.
XX
PI   Millis AJT;
XX
DR   WPI; 2000-431300/37.
XX
PT   Clusterin and gp38K-related peptide capable of altering cell migration
PT   useful for treating atherosclerosis, cancer and stenosis following
PT   vascular trauma or disease  -
XX
PS   Disclosure; Fig 2; 43pp; English.
XX
CC   The present sequence is the protein sequence of gp38k. Gp38k, a
CC   chemoattractant, is essential for the migration of vascular smooth muscle
CC   cells (VSMC). The gene and protein can, therefore, be used to promote
CC   wound healing, angiogenesis and vasculogenesis, in the treatment of
CC   stenosis following vascular trauma or disease and to treat
CC   atherosclerosis, and antisense sequences can be used to treat cancer, as
CC   angiogenesis is vital for tumour survival.
XX
SQ   Sequence   383 AA;
```

```
  Query Match           100.0%;  Score 66;  DB 21;  Length 383;
  Best Local Similarity  100.0%;  Pred. No. 0.00045;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SQRFSKIASNTQSR 14
              ||||||||||||||
Db        103 SQRFSKIASNTQSR 116


RESULT 12
ABG96297
ID   ABG96297 standard; Protein; 383 AA.
XX
AC   ABG96297;
XX
DT   11-DEC-2002  (first entry)
XX
DE   Human ovarian cancer marker OV6.
XX
KW   Human; ovarian cancer; marker; cancer; familial history; brain disorder;
KW   central nervous system disorder; bacterial meningitis; viral meningitis;
KW   Alzheimer's disease; Parkinson's disease; cerebral oedema; hydrocephalus;
KW   brain herniation; inflammation; encephalitis; testicular disorder;
KW   nontuberculous granulomatous orchitis; connective tissue disorder;
KW   heart disorder; ischaemic heart disease; atherosclerosis; neoplasm;
KW   histological type; carcinogenic; ovarian cancer marker.
XX
OS   Homo sapiens.
XX
PN   WO200271928-A2.
XX
PD   19-SEP-2002.
XX
PF   14-MAR-2002; 2002WO-US07826.
XX
PR   14-MAR-2001; 2001US-276025P.
PR   14-MAR-2001; 2001US-276026P.
PR   10-AUG-2001; 2001US-311732P.
PR   19-SEP-2001; 2001US-323580P.
PR   26-SEP-2001; 2001US-324967P.
PR   26-SEP-2001; 2001US-325102P.
PR   26-SEP-2001; 2001US-325149P.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   Monahan JE,  Gannavarapu M,  Hoersch S,  Kamalkar S,  Kovatis SG;
PI   Meyers RE,  Morrisey MP,  Olandt PJ,  Sen A,  Vieby PD,  Mills GB;
PI   Bast RC,  Lu K,  Schmandt RE,  Zhao X,  Glatt K;
XX
DR   WPI; 2002-723277/78.
DR   N-PSDB; ABS76389.
XX
PT   Assessing whether a patient is afflicted with ovarian cancer, useful in
PT   assessing the stage or progression of the disease, comprises comparing
PT   the expression level of a cancer marker in a sample from a patient and
PT   from a non cancer patient -
XX
PS   Disclosure; Page 164-165; 481pp; English.
XX
CC   The present invention relates to a new method for assessing whether a
CC   patient is afflicted with ovarian cancer. The method involves comparing
CC   the expression level of a marker in a patient sample and the normal level
CC   of expression of the marker in a control non-ovarian cancer sample, where
CC   the marker is selected from 363 cancer markers described in the
CC   specification. The method of the invention is useful in diagnosing or
CC   characterising cancer, in detecting the presence of cancer as early as
CC   possible, and the recurrence of ovarian cancer. The method may also be of
CC   particular use with patients having an enhanced risk of developing
CC   ovarian cancer (e.g. patients having a familial history of ovarian
CC   cancer). The cancer markers may be used in the management and treatment
CC   of e.g. brain and central nervous system disorders (e.g. bacterial and
CC   viral meningitis, Alzheimer's disease or Parkinson's disease), brain
CC   disorders (e.g. cerebral oedema, hydrocephalus or brain herniations),
CC   inflammations (e.g. bacterial or viral meningitis or encephalitis),
CC   testicular disorders (e.g. nontuberculous granulomatous orchitis),
CC   connective tissue disorders, or heart disorders (e.g. ischaemic heart
CC   disease or atherosclerosis). The compositions and methods may also be
CC   used in assessing the histological type of neoplasm associated with
CC   ovarian cancer, monitoring the progression of ovarian cancer,
CC   determining whether ovarian cancer has metastasized or is likely to
CC   metastasize, selecting a composition for inhibiting ovarian cancer,
CC   assessing the ovarian carcinogenic potential of a compound, or
CC   inhibiting ovarian cancer or at risk of developing ovarian cancer. The
CC   present amino acid sequence represents one of the ovarian cancer markers
CC   described in the invention.
XX
SQ   Sequence   383 AA;

  Query Match           100.0%;  Score 66;  DB 23;  Length 383;
  Best Local Similarity  100.0%;  Pred. No. 0.00045;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SQRFSKIASNTQSR 14
              ||||||||||||||
Db        103 SQRFSKIASNTQSR 116


RESULT 13
ABU56651
ID   ABU56651 standard; Protein; 383 AA.
XX
AC   ABU56651;
XX
DT   02-APR-2003  (first entry)
XX
DE   Lung cancer-associated polypeptide #244.
XX
KW   Lung cancer-associated polypeptide; cytostatic; emphysema;
KW   antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
KW   small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
KW   chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
KW   interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
XX
OS   Unidentified.
XX
PN   WO2002086443-A2.
XX
PD   31-OCT-2002.
XX
PF   18-APR-2002; 2002WO-US12476.
XX
PR   18-APR-2001; 2001US-284770P.
PR   10-MAY-2001; 2001US-290492P.
PR   09-NOV-2001; 2001US-339245P.
PR   13-NOV-2001; 2001US-350666P.
PR   29-NOV-2001; 2001US-334370P.
PR   12-APR-2002; 2002US-372246P.
XX
PA   (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI   Aziz N,  Murray R;
XX
DR   WPI; 2003-093161/08.
DR   N-PSDB; ABX76380.
XX
PT   Detecting a lung cancer-associated transcript in a cell from a patient
PT   for treating lung cancer, by contacting a biological sample from the
PT   patient with a polynucleotide that exhibits increased or decreased
PT   expression in lung cancer  -
XX
PS   Claim 27; Page 377; 453pp; English.
XX
CC   The invention relates to a method for detecting a lung cancer-associated
CC   transcript in a cell from a patient, comprising contacting a biological
CC   sample from the patient with a polynucleotide that selectively hybridises
CC   to a sequence that is at least 80 % identical to a gene that exhibits
```

```
CC   increased or decreased expression in lung cancer samples. Lung
CC   cancer-associated polynucleotides and polypeptides are used for
CC   identifying a compound that modulates a lung cancer-associated
CC   polypeptide, for inhibiting proliferation of a lung cancer-associated
CC   cell to treat lung cancer in a patient and for treating a mammal having
CC   lung cancer by administering a modulatory compound identified. The
CC   methods are useful for treating lung cancer, such as small cell lung
CC   cancer, non-small cell lung cancer or other benign or precancerous
CC   lesions, e.g. atelectasis, emphysema, bronchitis, chronic obstructive
CC   pulmonary disease, fibrosis, hypersensitivity pneumonitis, interstitial
CC   pulmonary fibrosis, asthma and bronchiectasis. The genes, polynucleotides
CC   and polypeptides are useful for diagnostic purposes and as targets for
CC   screening for therapeutic compounds that modulate lung cancer, such as
CC   antibodies. Sequences ABU56408-ABU56745 represent lung cancer-associated
CC   polypeptides of the invention.
XX
SQ   Sequence   383 AA;

  Query Match             100.0%;  Score 66;  DB 24;  Length 383;
  Best Local Similarity   100.0%;  Pred. No. 0.00045;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SQRFSKIASNTQSR 14
              ||||||||||||||
Db        103 SQRFSKIASNTQSR 116


RESULT 14
AAU02077
ID   AAU02077 standard; Protein; 459 AA.
XX
AC   AAU02077;
XX
DT   07-SEP-2001  (first entry)
XX
DE   Synthetic multi-target autoantigen Y-RAPh.
XX
KW   Y-RAPa; immunogenic epitope cluster; IEC;
KW   autoantigen; shCollagen/RA; shAggrecan/RA; shGLP/RA; shHSP/RA;
KW   autoimmune disease; multiple sclerosis; rheumatoid arthritis;
KW   insulin-dependent diabetes mellitus; myasthenia gravis; uveitis;
KW   autoimmune hepatitis; thyroiditis; insulitis; orchitis;
KW   idiopathic thrombocytopenic purpura; inflammatory disease;
KW   Crohn's disease; ulcerative colitis; shMultiTAG.
XX
OS   Synthetic.
OS   Homo sapiens.
OS   Escherichia coli.
OS   Mycobacterium tuberculosis.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 455
FT                   /label= OTHER
FT                   /note= "Other= STOP"
FT   Misc-difference 457
FT                   /label= OTHER
FT                   /note= "Other= STOP"
XX
PN   WO200131037-A2.
XX
PD   03-MAY-2001.
XX
PF   26-OCT-2000; 2000WO-IL00688.
XX
PR   27-OCT-1999;   99IL-0132611.
XX
PA   (YEDA ) YEDA RES & DEV CO LTD.
XX
PI   Ben-Nun A,  Kerlero De Rosbo N,  Sappler GP;
XX
DR   WPI; 2001-300515/31.
DR   N-PSDB; AAS04838.
XX
PT   Novel synthetic human target autoantigen gene useful for treating
PT   autoimmune diseases such as multiple sclerosis, insulin-dependent
PT   diabetes mellitus, rheumatoid arthritis, myasthenia gravis, and uveitis
PT   -
XX
PS   Claim 182; Fig 57; 182pp; English.
XX
CC   The sequence represents synthetic human multi-target autoantigen
CC   (shMultiTAG) Y-RAPa consisting of regions encoding shCollagen/RA,
CC   shAggrecan/RA, shGLP/RA and shHSP/RA. The synthetic human target
CC   autoantigen genes of the invention comprise sequences coding for at least
CC   2 IECs of autoantigen(s) related to a specific autoimmune disease. The
CC   synthetic human target autoantigen genes are useful for treating
CC   autoimmune diseases such as multiple sclerosis, insulin-dependent
CC   diabetes mellitus, rheumatoid arthritis, myasthenia gravis, uveitis,
CC   autoimmune hepatitis, thyroiditis, insulitis, orchitis, idiopathic
CC   thrombocytopenic purpura, and inflammatory diseases (Crohn's disease,
CC   ulcerative colitis). The synthetic human target autoantigen genes are
CC   also useful for diagnosis and/or monitoring the progression of the
CC   autoimmune disease.
XX
SQ   Sequence   459 AA;

  Query Match             100.0%;  Score 66;  DB 22;  Length 459;
  Best Local Similarity   100.0%;  Pred. No. 0.00055;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SQRFSKIASNTQSR 14
              ||||||||||||||
Db        253 SQRFSKIASNTQSR 266


RESULT 15
AAW45113
ID   AAW45113 standard; peptide; 16 AA.
XX
AC   AAW45113;
XX
DT   28-APR-1998  (first entry)
XX
DE   Human cartilage glycoprotein 39 derived peptide #14.
XX
KW   Articular cartilage; immunosuppressive therapy; antigen; autoantigen;
KW   immunological tolerance; T-cell; human cartilage glycoprotein 39;
KW   HC gp-39; rheumatoid arthritis; epitope.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9740068-A1.
XX
PD   30-OCT-1997.
XX
PF   22-APR-1997;   97WO-EP02051.
XX
PR   24-APR-1996;   96EP-0201106.
XX
PA   (ALKU ) AKZO NOBEL NV.
XX
PI   Boots AMH,  Verheijden GFM;
XX
DR   WPI; 1997-535775/49.
XX
PT   Peptide suitable for use in antigen specific immunosuppressive
PT   therapy - resembles or mimics epitope present on HC gp-39, so
PT   inducing systemic immunological tolerance to rheumatoid arthritis
PT   autoantigen
XX
PS   Disclosure; Page 15; 82pp; English.
XX
CC   The present sequence represents a peptide which resembles or mimics an
CC   epitope present on human cartilage glycoprotein 39 (HC gp-39), an
CC   autoantigen in rheumatoid arthritis. The invention relates to peptides
```

CC consisting of 16-55 amino acid residues comprising at least one of the
CC following 19 sequences: LVCYYTSWS; PLCTHIIYS; IIYSFANIS; LKTLLSVGG;
CC PIKSVPPPL; FDGLDLAWL; LYPGRRDKQ; YDIAKISQH; LDFISIMTY; FISIMIYDF;
CC FRGQEDASP; YAVGYMLRL; MLRLGAPAS; LAYYEICDF; LRGATVHRI; YLKDRQLAG;
CC LAGAMVWAL; VWALDLDDF; or LDIDDFQGS. They can be used medically in
CC antigen specific immunosuppressive therapy, particularly the treatment
CC of T-cell mediated destruction of articular cartilage in autoimmune
CC diseases (e.g. rheumatoid arthritis). They can also be used to detect
CC activated autoreactive T cells in an individual. The peptides have a
CC specific effect on the autoreactive T cells, thus leaving the other
CC components of the immune system intact, unlike the non-specific
CC suppressive effect of immunosuppressive drugs, and do not cause toxic
CC side effects. The peptides are predominantly recognised by autoreactive
CC T cells from rheumatoid arthritis patients, but rarely by those from
CC healthy donors.
XX
SQ Sequence 16 AA;

Query Match 92.4%; Score 61; DB 16; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.00011;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SQRFSKIASNTQS 13
|||||||||||||
Db 4 SQRFSKIASNTQS 16

Search completed: September 24, 2003, 15:41:24
Job time : 57.2823 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 24, 2003, 13:35:23 ; Search time 16.9888 Seconds
(without alignments)
34.867 Million cell updates/sec

Title: US-09-744-282-8
Perfect score: 66
Sequence: 1 SQRFSKIASNTQSR 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310856 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	66	100.0	14	1	US-08-619-645-8	Sequence 8, Appli
2	66	100.0	14	2	US-08-634-493-8	Sequence 8, Appli
3	66	100.0	383	4	US-09-459-743D-17	Sequence 17, Appl
4	61	92.4	16	3	US-09-171-705-14	Sequence 14, Appl
5	52	78.8	16	3	US-09-171-705-15	Sequence 15, Appl
6	43	65.2	9	1	US-08-619-645-4	Sequence 4, Appli
7	43	65.2	9	2	US-08-634-493-4	Sequence 4, Appli
8	37	56.1	239	4	US-09-252-991A-20802	Sequence 20802, A
9	34	51.5	480	1	US-07-303-638A-2	Sequence 2, Appli
10	33	50.0	16	3	US-09-171-705-13	Sequence 13, Appl
11	33	50.0	340	1	US-08-307-409-32	Sequence 32, Appl
12	33	50.0	340	3	US-09-209-268-32	Sequence 32, Appl
13	33	50.0	390	4	US-09-198-452A-254	Sequence 254, App
14	33	50.0	453	4	US-09-252-991A-23266	Sequence 23266, A
15	33	50.0	682	4	US-09-252-991A-23129	Sequence 23129, A
16	33	50.0	738	3	US-08-478-208-32	Sequence 32, Appl
17	33	50.0	738	4	US-09-336-536-73	Sequence 73, Appl
18	33	50.0	738	6	5264554-2	Patent No. 5264554
19	33	50.0	1778	4	US-09-252-991A-18159	Sequence 18159, A
20	33	50.0	3169	2	US-08-477-451-6	Sequence 6, Appli
21	32	48.5	81	4	US-09-198-452A-1255	Sequence 1255, Ap
22	32	48.5	95	4	US-09-107-532A-5190	Sequence 5190, Ap
23	32	48.5	102	4	US-09-134-001C-3804	Sequence 3804, Ap
24	32	48.5	138	4	US-09-134-001C-4796	Sequence 4796, Ap
25	32	48.5	179	4	US-09-732-210-70	Sequence 70, Appl
26	32	48.5	486	4	US-09-328-352-6315	Sequence 6315, Ap
27	32	48.5	610	4	US-09-252-991A-20409	Sequence 20409, A
28	32	48.5	1046	4	US-09-252-991A-27508	Sequence 27508, A
29	31	47.0	157	4	US-09-252-991A-25419	Sequence 25419, A
30	31	47.0	191	3	US-09-198-484-5	Sequence 5, Appli
31	31	47.0	245	4	US-09-107-532A-4448	Sequence 4448, Ap
32	31	47.0	267	4	US-09-252-991A-17521	Sequence 17521, A
33	31	47.0	316	4	US-09-504-358-10	Sequence 10, Appl
34	31	47.0	316	4	US-09-954-314-10	Sequence 10, Appl
35	31	47.0	419	4	US-09-328-352-6451	Sequence 6451, Ap
36	31	47.0	431	4	US-09-600-370-1	Sequence 1, Appli
37	31	47.0	495	3	US-09-198-484-4	Sequence 4, Appli
38	31	47.0	580	4	US-09-134-001C-5611	Sequence 5611, Ap
39	31	47.0	596	4	US-09-252-991A-21255	Sequence 21255, A
40	31	47.0	682	3	US-08-613-009A-10	Sequence 10, Appl
41	31	47.0	682	4	US-08-775-570B-12	Sequence 12, Appl
42	31	47.0	682	4	US-09-059-584-12	Sequence 12, Appl
43	31	47.0	702	2	US-08-567-941-25	Sequence 25, Appl
44	31	47.0	702	3	US-08-613-003A-9	Sequence 9, Appli
45	31	47.0	702	3	US-09-074-658-25	Sequence 25, Appl

ALIGNMENTS

RESULT 1
US-08-619-645-8
; Sequence 8, Application US/08619645
; Patent No. 5736507
; GENERAL INFORMATION:
; APPLICANT: Boots, Anna M.H.
; APPLICANT: Verheijden, Gijsbertus F.M.
; TITLE OF INVENTION: No. 5736507el peptides derived from
; TITLE OF INVENTION: autoantigen for use in immunotherapy of autoimmune
; TITLE OF INVENTION: disease
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Akzo No. 5736507el Patent Department
; STREET: 1300 Piccard Drive, Suite 206
; CITY: Rockville
; STATE: Maryland
; COUNTRY: U.S.A.
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25(EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/619,645
; FILING DATE: 25-MAR-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Gormley, Mary E.
; REGISTRATION NUMBER: 34,409
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 947-4433
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-619-645-8

Query Match 100.0%; Score 66; DB 1; Length 14;
Best Local Similarity 100.0%; Pred. No. 4.5e-06;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SQRFSKIASNTQSR 14
||||||||||||||
Db 1 SQRFSKIASNTQSR 14

```
RESULT 2
US-08-634-493-8
; Sequence 8, Application US/08634493
; Patent No. 5843449
;  GENERAL INFORMATION:
;    APPLICANT:  A.M.H. Boots
;    APPLICANT:  G.F.M. Verheijden
;    APPLICANT:  E.S. Bos
;    TITLE OF INVENTION:  No. 5843449el Peptides derived from autoantigen for use
;    TITLE OF INVENTION:  in Immunotherapy of Autoimmune Diseases
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Akzo No. 5843449el Patent Department
;      STREET:  1300 Piccard Drive, Suite 206
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/634,493
;      FILING DATE:  18-APR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/619,845
;      FILING DATE:  25-MAR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/EP95/04201
;      FILING DATE:  25-OCT-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  NL 9402031287
;      FILING DATE:  27-OCT-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  NL 9520008860
;      FILING DATE:  07-APR-1995
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 258-5200
;      TELEFAX:  (301) 977-0847
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  14 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-634-493-8

  Query Match             100.0%;  Score 66;  DB 2;  Length 14;
  Best Local Similarity   100.0%;  Pred. No. 4.5e-06;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SQRFSKIASNTQSR 14
              ||||||||||||||
Db          1 SQRFSKIASNTQSR 14


RESULT 3
US-09-459-749D-17
; Sequence 17, Application US/09459749D
; Patent No. 6464975
; GENERAL INFORMATION:
;  APPLICANT: Millis, Albert J. T.
;  TITLE OF INVENTION: Compositions and Methods For Altering Cell Migration
;  FILE REFERENCE: 0794.016A
;  CURRENT APPLICATION NUMBER: US/09/459,749D
;  CURRENT FILING DATE:  1999-12-10
;  PRIOR APPLICATION NUMBER: 60/111,856
;  PRIOR FILING DATE: 1998-12-11
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
;    LENGTH: 383
;    TYPE: PRT
;    ORGANISM: Sus scrofa
US-09-459-749D-17

  Query Match             100.0%;  Score 66;  DB 4;  Length 383;
  Best Local Similarity   100.0%;  Pred. No. 0.00016;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SQRFSKIASNTQSR 14
              ||||||||||||||
Db        103 SQRFSKIASNTQSR 116


RESULT 4
US-09-171-705-14
; Sequence 14, Application US/09171705
; Patent No. 6184204
; GENERAL INFORMATION:
;  APPLICANT: BOOTS, ANNA M.H.
;  APPLICANT: VERHEIJDEN, GILBERTUS F.M.
;  TITLE OF INVENTION: NOVEL PEPTIDES SUITABLE FOR USE IN ANTIGEN SPECIFIC
;  TITLE OF INVENTION: IMMUNOSUPPRESSIVE THERAPY
;  FILE REFERENCE: O/96198 US
;  CURRENT APPLICATION NUMBER: US/09/171,705
;  CURRENT FILING DATE: 1999-02-09
;  NUMBER OF SEQ ID NOS: 78
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
;    LENGTH: 16
;    TYPE: PRT
;    ORGANISM: Artificial Sequence
;    FEATURE:
;    OTHER INFORMATION: Description of Artificial Sequence: DERIVED FROM
;    OTHER INFORMATION: SEQUENCE OF HUMAN CARTILAGE (HC) -39 PROTEIN
US-09-171-705-14

  Query Match             92.4%;  Score 61;  DB 3;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 4.4e-05;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SQRFSKIASNTQS 13
              |||||||||||||
Db          4 SQRFSKIASNTQS 16


RESULT 5
US-09-171-705-15
; Sequence 15, Application US/09171705
; Patent No. 6184204
; GENERAL INFORMATION:
;  APPLICANT: BOOTS, ANNA M.H.
;  APPLICANT: VERHEIJDEN, GILBERTUS F.M.
;  TITLE OF INVENTION: NOVEL PEPTIDES SUITABLE FOR USE IN ANTIGEN SPECIFIC
;  TITLE OF INVENTION: IMMUNOSUPPRESSIVE THERAPY
;  FILE REFERENCE: O/96198 US
;  CURRENT APPLICATION NUMBER: US/09/171,705
;  CURRENT FILING DATE: 1999-02-09
;  NUMBER OF SEQ ID NOS: 78
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
;    LENGTH: 16
;    TYPE: PRT
;    ORGANISM: Artificial Sequence
;    FEATURE:
;    OTHER INFORMATION: Description of Artificial Sequence: DERIVED FROM
;    OTHER INFORMATION: SEQUENCE OF HUMAN CARTILAGE (HC) -39 PROTEIN
US-09-171-705-15

  Query Match             78.8%;  Score 52;  DB 3;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 0.002;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          4 FSKIASNTQSR 14
              ||:|.||||:'
Db          1 FSKIASNTQSR 11


RESULT 6
US-08-619-645-4
; Sequence 4, Application US/08619645
; Patent No. 5736507
;  GENERAL INFORMATION:
;    APPLICANT:  Boots, Anna M.H.
;    APPLICANT:  Verheijden, Gijsbertus F.M.
;    TITLE OF INVENTION:  No. 5736507el peptides derived from
;    TITLE OF INVENTION:  autoantigen for use in immunotherapy of autoimmune
;    TITLE OF INVENTION:  disease
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Akzo No. 5736507el Patent Department
;      STREET:  1300 Piccard Drive, Suite 206
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  U.S.A.
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25(EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/619,645
;      FILING DATE:  25-MAR-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Gormley, Mary E.
;      REGISTRATION NUMBER:  34,409
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 947-4433
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-619-645-4

  Query Match             65.2%;  Score 43;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 FSKIASNTQ 12
              ||:|.|||'
Db          1 FSKIASNTQ 9


RESULT 7
US-08-634-493-4
; Sequence 4, Application US/08634493
; Patent No. 5843449
;  GENERAL INFORMATION:
;    APPLICANT:  A.M.H. Boots
;    APPLICANT:  G.F.M. Verheijden
;    APPLICANT:  E.S. Bos
;    TITLE OF INVENTION:  No. 5843449el Peptides derived from autoantigen for use
;    TITLE OF INVENTION:  in Immunotherapy of Autoimmune Diseases
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Akzo No. 5843449el Patent Department
;      STREET:  1300 Piccard Drive, Suite 206
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/634,493
;      FILING DATE:  18-APR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/619,645
;      FILING DATE:  25-MAR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/EP95/04201
;      FILING DATE:  25-OCT-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  NL 9420l287
;      FILING DATE:  27-OCT-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  NL 952008860
;      FILING DATE:  07-APR-1995
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 258-5200
;      TELEFAX:  (301) 977-0947
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-634-493-4

  Query Match             65.2%;  Score 43;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 FSKIASNTQ 12
              ||:|.|||'
Db          1 FSKIASNTQ 9


RESULT 8
US-09-252-991A-20802
; Sequence 20802, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMON
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER:  US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/054,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 20802
;   LENGTH:  233
;   TYPE:  PRT
;   ORGANISM:  Pseudomonas aeruginosa
US-09-252-991A-20802

  Query Match             56.1%;  Score 37;  DB 4;  Length 233;
  Best Local Similarity   50.0%;  Pred. No. 22;
  Matches    7;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 SQRFSKIASNTQSR 14
              :: ||:|:||: ||
Db         87 ARAFSRISSNSASR 100
```

```
RESULT 9
US-07-803-636A-2
; Sequence 2, Application US/07803636A
; Patent No. 5422428
;  GENERAL INFORMATION:
;    APPLICANT:  McGUIRE, TRAVIS C., TERRY F. McELWAIN, LANCE E. PERRYMAN,
;    APPLICANT:  WILLIAM C. DAVIS
;    TITLE OF INVENTION:  IMMUNIZATION AGAINST BABESIOSIS USING
;    TITLE OF INVENTION:  PURIFIED SURFACE ANTIGENS OF BABESIA BIGEMINA AND SIMILAR
;    TITLE OF INVENTION:  IMMUNOGENS
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  DAVID R. SALIWANCHIK
;      STREET:  2421 NW 41ST STREET, SUITE A-1
;      CITY:  GAINESVILLE
;      STATE:  FLORIDA
;      COUNTRY:  USA
;      ZIP:  32606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/803,636A
;      FILING DATE:  13911206
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  SALIWANCHIK, DAVID R
;      REGISTRATION NUMBER:  31,794
;      REFERENCE/DOCKET NUMBER:  WA4-059.C1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  904-375-8100
;      TELEFAX:  904-372-5800
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  480 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-07-803-636A-2

  Query Match             51.5%;  Score 34;  DB 1;  Length 480;
  Best Local Similarity   54.5%;  Pred. No. 1.7e-02;
  Matches    6;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 SQRFSKIASNT 11
              ::|:|:|:|: |
Db        266 AERFSKMATKT 276


RESULT 10
US-09-171-705-13
; Sequence 13, Application US/09171705
; Patent No. 6184204
; GENERAL INFORMATION:
;  APPLICANT: BOOTS, ANNA M.H.
;  APPLICANT: VERHEIJDEN, GILBERTUS F.M.
;  TITLE OF INVENTION: NOVEL PEPTIDES SUITABLE FOR USE IN ANTIGEN SPECIFIC
;  TITLE OF INVENTION: IMMUNOSUPPRESSIVE THERAPY
;  FILE REFERENCE: C/96198 US
;  CURRENT APPLICATION NUMBER: US/09/171,705
;  CURRENT FILING DATE: 1999-02-09
;  NUMBER OF SEQ ID NOS: 78
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
;   LENGTH: 16
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: DERIVED FROM
;   OTHER INFORMATION: SEQUENCE OF HUMAN CARTILAGE (HC) -39 PROTEIN
US-09-171-705-13

  Query Match             50.0%;  Score 33;  DB 3;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 6.7;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SQRFSKI 7
              |||||||
Db         10 SQRFSKI 16


RESULT 11
US-08-307-499-32
; Sequence 32, Application US/08307499
; Patent No. 5651972
;  GENERAL INFORMATION:
;    APPLICANT:  Moyer, Richard W.
;    APPLICANT:  Vi uela, Eladio
;    APPLICANT:  Gibbs, E.P.J.
;    TITLE OF INVENTION:  Use of Recombinant Swine Poxvirus as a
;    TITLE OF INVENTION:  Live Vaccine Vector
;    NUMBER OF SEQUENCES:  60
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  David R. Saliwanchik
;      STREET:  2421 N.W. 41st Street, Suite A-1
;      CITY:  Gainesville
;      STATE:  Florida
;      COUNTRY:  U.S.A.
;      ZIP:  32606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/307,499
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/908,241
;      FILING DATE:  1-JUL-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/908,630
;      FILING DATE:  29-JUN-1992
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/842,212
;      FILING DATE:  21-APR-1992
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Saliwanchik, David R.
;      REGISTRATION NUMBER:  31,794
;      REFERENCE/DOCKET NUMBER:  UF35.1.PWCC1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  904-375-8100
;      TELEFAX:  904-372-5800
;  INFORMATION FOR SEQ ID NO:  32:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  340 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-307-499-32

  Query Match             50.0%;  Score 33;  DB 1;  Length 340;
  Best Local Similarity   54.5%;  Pred. No. 1.8e+02;
  Matches    6;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 SQRFSKIASNT 11
              ::|:|:|: | |
Db        147 NKRFSKVRSRT 157


RESULT 12
```

```
US-09-299-268-32
; Sequence 32, Application US/09299268
; Patent No. 6217882
;  GENERAL INFORMATION:
;    APPLICANT:  Moyer, Richard W.
;    APPLICANT:  Vi uela, Eladio
;    APPLICANT:  Gibbs, E.P.J.
;    TITLE OF INVENTION:  Use of Recombinant Swine Poxvirus as a
;    TITLE OF INVENTION:  Live Vaccine Vector
;    NUMBER OF SEQUENCES:  60
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  David R. Saliwanchik
;      STREET:  2421 N.W. 41st Street, Suite A-1
;      CITY:  Gainesville
;      STATE:  Florida
;      COUNTRY:  U.S.A.
;      ZIP:  32606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  Patentin Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/299,268
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/901,127
;      FILING DATE:
;      APPLICATION NUMBER:  US 07/908,241
;      FILING DATE:  1-JUL-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/908,630
;      FILING DATE:  29-JUN-1992
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/342,212
;      FILING DATE:  21-APR-1992
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Saliwanchik, David R.
;      REGISTRATION NUMBER:  31,794
;      REFERENCE/DOCKET NUMBER:  UF35.1.FWC01
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  904-375-8100
;      TELEFAX:  904-372-5800
;  INFORMATION FOR SEQ ID NO:  32:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  340 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-299-268-32

  Query Match             50.0%;  Score 33;  DB 3;  Length 340;
  Best Local Similarity   54.5%;  Pred. No. 1.8e+02;
  Matches    6;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 SQRFSKIASNT 11
              :: |:|: ' :
Db        147 NKRFSKVRSRT 157


RESULT 13
US-09-198-452A-254
; Sequence 254, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
;  APPLICANT: Griffais, R.
;  TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
;  TITLE OF INVENTION:  thereof and uses thereof, in particular for the diagnosis, prevo
;  TITLE OF INVENTION:  and treatment of infection
;  FILE REFERENCE: 9710-003-999
;  CURRENT APPLICATION NUMBER: US/09/198,452A
;  CURRENT FILING DATE:  1998-11-24
;  NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 254
;   LENGTH: 390
;   TYPE: PRT
;   ORGANISM: Chlamydia pneumoniae
US-09-198-452A-254

  Query Match             50.0%;  Score 33;  DB 4;  Length 390;
  Best Local Similarity   57.1%;  Pred. No. 2.1e+02;
  Matches    6;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 SQRFSKIASNTQSR 14
              |   : |||:|: |
Db         15 SSHEASIASNTQVR 28


RESULT 14
US-09-252-991A-23266
; Sequence 23266, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONA
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 23266
;   LENGTH: 453
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: (389),(443)
;   OTHER INFORMATION: Identity of amino acid at the above locations are unknown.
US-09-252-991A-23266

  Query Match             50.0%;  Score 33;  DB 4;  Length 453;
  Best Local Similarity   66.7%;  Pred. No. 2.5e-02;
  Matches    6;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 QRFSKIASN 10
              :|||:|:|:|
Db         35 RRFSRIAAN 43


RESULT 15
US-09-252-991A-23129
; Sequence 23129, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONA
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 23129
;   LENGTH: 632
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
```

US-09-252-991A-23129

Query Match 50.0%; Score 33; DB 4; Length 632;
Best Local Similarity 42.9%; Pred. No. 3.6e+02;
Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

```
Qy          1 SQRFSKIASNTQSR 14
              |:|||: |:  : |
Db        376 SERFSRAAARARQR 389
```

Search completed: September 24, 2003, 13:51:17
Job time : 17.9888 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 24, 2003, 13:47:40 ; Search time 31.3034 Seconds
(without alignments)
87.670 Million cell updates/sec

Title: US-09-744-282-8
Perfect score: 66
Sequence: 1 SQRFSKIASNTQSR 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 566894 seqs, 151307093 residues

Total number of hits satisfying chosen parameters: 566894

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	66	100.0	383	10	US-09-459-749D-17	Sequence 17, Appl
2	66	100.0	383	15	US-10-097-340-45	Sequence 45, Appl
3	40	60.6	794	15	US-10-128-714-8165	Sequence 8165, Ap
4	40	60.6	1138	15	US-10-128-714-8165	Sequence 8165, Ap
5	36	54.5	78	15	US-10-106-698-5185	Sequence 5185, Ap
6	35	53.0	371	10	US-09-923-844B-2	Sequence 2, Appli
7	35	53.0	1230	10	US-09-881-752A-150	Sequence 150, App
8	34	51.5	180	12	US-09-896-923-40	Sequence 40, Appl
9	34	51.5	407	15	US-10-169-048-30	Sequence 30, Appl
10	34	51.5	704	15	US-10-213-990-21	Sequence 21, Appl
11	34	51.5	867	10	US-09-801-368-98	Sequence 98, Appl
12	34	51.5	1042	15	US-10-157-031-108	Sequence 108, App
13	34	51.5	1516	15	US-10-156-761-13716	Sequence 13716, A
14	34	51.5	1981	10	US-09-928-457-38	Sequence 38, Appl
15	34	51.5	2015	15	US-10-066-551-1	Sequence 1, Appli
16	33	50.0	120	15	US-10-106-698-4440	Sequence 4440, Ap
17	33	50.0	267	15	US-10-156-761-12545	Sequence 12545, A
18	33	50.0	340	10	US-09-976-605-8	Sequence 8, Appli
19	33	50.0	477	15	US-10-204-887-104	Sequence 104, App
20	33	50.0	549	9	US-09-925-302-477	Sequence 477, App
21	33	50.0	656	12	US-10-242-576-10	Sequence 10, Appl
22	33	50.0	656	12	US-10-242-576-12	Sequence 12, Appl
23	33	50.0	1308	10	US-09-862-927-79	Sequence 79, Appl
24	32	48.5	46	9	US-09-764-860-509	Sequence 509, App
25	32	48.5	46	15	US-10-074-095-509	Sequence 509, App
26	32	48.5	191	15	US-10-286-264-126	Sequence 126, App
27	32	48.5	192	15	US-10-021-811-50	Sequence 50, Appl
28	32	48.5	256	9	US-09-867-550-684	Sequence 684, App
29	32	48.5	257	16	US-10-080-170-217	Sequence 217, App
30	32	48.5	286	10	US-09-738-626-5772	Sequence 5772, Ap
31	32	48.5	324	9	US-09-869-562-40	Sequence 40, Appl
32	32	48.5	324	12	US-10-303-664A-39	Sequence 39, Appl
33	32	48.5	381	10	US-09-895-913A-88	Sequence 88, Appl
34	32	48.5	733	12	US 10-291-583-61	Sequence 61, Appl
35	32	48.5	733	12	US-10-291-583-62	Sequence 62, Appl
36	32	48.5	736	12	US 10-291-583-65	Sequence 65, Appl
37	32	48.5	779	9	US-09-864-761-46668	Sequence 46668, A
38	32	48.5	913	9	US-09-801-574-65	Sequence 65, Appl
39	31	47.0	108	10	US-09-881-752A-236	Sequence 236, App
40	31	47.0	121	9	US-09-864-761-48435	Sequence 48435, A
41	31	47.0	124	16	US-10-080-170-523	Sequence 523, App
42	31	47.0	140	15	US-10-080-980-5	Sequence 5, Appli
43	31	47.0	150	14	US-10-001-857-203	Sequence 203, App
44	31	47.0	179	10	US-09-796-692-2177	Sequence 2177, Ap
45	31	47.0	179	15	US-10-040-862-2177	Sequence 2177, Ap

ALIGNMENTS

RESULT 1
US-09-459-749D-17
; Sequence 17, Application US/09459749D
; Patent No. US20020136716A1
; GENERAL INFORMATION:
; APPLICANT: Millis, Albert J. T.
; TITLE OF INVENTION: Compositions and Methods For Altering Cell Migration
; FILE REFERENCE: 0794.016A
; CURRENT APPLICATION NUMBER: US/09/459,749D
; CURRENT FILING DATE: 1999-12-10
; PRIOR APPLICATION NUMBER: 60/111,856
; PRIOR FILING DATE: 1998-12-11
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 383
; TYPE: PRT
; ORGANISM: Sus scrofa
US-09-459-749D-17

Query Match 100.0%; Score 66; DB 10; Length 383;
Best Local Similarity 100.0%; Pred. No. 0.00036;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SQRFSKIASNTQSR 14
||||||||||||||
Db 103 SQRFSKIASNTQSR 116

RESULT 2
US-10-097-340-45
; Sequence 45, Application US/10097340
; Publication No. US20030087250A1
; GENERAL INFORMATION:
; APPLICANT: John MONAHAN
; APPLICANT: Manjula GANNAVARAPU
; APPLICANT: Sebastian HOERSCH
; APPLICANT: Shubhangi KAMATKAR

```
;  APPLICANT:  Steve G. KOVATS
;  APPLICANT:  Rachel F. MEYERS
;  APPLICANT:  Michael MORRISEY
;  APPLICANT:  Peter OLANDT
;  APPLICANT:  Ami SEN
;  APPLICANT:  Peter VEIBY
;  APPLICANT:  Gordon B. MILLS
;  APPLICANT:  Robert C. BAST, Jr.
;  APPLICANT:  Karen LU
;  APPLICANT:  Rosemarie SCHMANDT
;  APPLICANT:  Xumei ZHAO
;  APPLICANT:  Karen GLAIT
;  TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
;  TITLE OF INVENTION:  Assessment, Prevention, and Therapy of Ovarian Cancer
;  FILE REFERENCE: MRI-030
;  CURRENT APPLICATION NUMBER: US/10/097,340
;  CURRENT FILING DATE:  2002-03-14
;  PRIOR APPLICATION NUMBER: 60/276,025
;  PRIOR FILING DATE: 2001-03-14
;  PRIOR APPLICATION NUMBER: 60/325,149
;  PRIOR FILING DATE: 2001-09-26
;  PRIOR APPLICATION NUMBER: 60/276,026
;  PRIOR FILING DATE: 2001-03-14
;  PRIOR APPLICATION NUMBER: 60/324,967
;  PRIOR FILING DATE: 2001/09/25
;  PRIOR APPLICATION NUMBER: 60/311,732
;  PRIOR FILING DATE: 2001-08-10
;  PRIOR APPLICATION NUMBER: 60/325,102
;  PRIOR FILING DATE: 2001-09-26
;  PRIOR APPLICATION NUMBER: 60/323,580
;  PRIOR FILING DATE: 2001-09-19
;  NUMBER OF SEQ ID NOS: 363
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 45
;   LENGTH: 383
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-097-340-45

  Query Match           100.0%;  Score 66;  DB 15;  Length 383;
  Best Local Similarity  100.0%;  Pred. No. 0.00036;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SQRFSKIASNTQSR 14
              ||||||||||||||
Db        103 SQRFSKIASNTQSR 116


RESULT 3
US-10-128-714-3165
; Sequence 3165, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Bo
;  APPLICANT:  Hu, Wenqi
;  APPLICANT:  Tishkoff, Daniel
;  APPLICANT:  Zamudio, Carlos
;  APPLICANT:  Eroshkin, Alexey M
;  APPLICANT:  Lemieux, Sebastien M
;  TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
;  TITLE OF INVENTION:  Methods of Use
;  FILE REFERENCE: 10182-018-999
;  CURRENT APPLICATION NUMBER: US/10/128,714
;  CURRENT FILING DATE:  2002-04-23
;  PRIOR APPLICATION NUMBER: US 60/285,697
;  PRIOR FILING DATE: 2001-04-23
;  PRIOR APPLICATION NUMBER: US 60/287,066
;  PRIOR FILING DATE: 2001-04-27
;  PRIOR APPLICATION NUMBER: US 60/295,890
;  PRIOR FILING DATE: 2001-06-05
;  PRIOR APPLICATION NUMBER: US 60/303,899
;  PRIOR FILING DATE: 2001-07-09
;  PRIOR APPLICATION NUMBER: US 60/316,362
;  PRIOR FILING DATE: 2001-08-31
;  NUMBER OF SEQ ID NOS: 8603
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3165
;   LENGTH: 794
;   TYPE: PRT
;   ORGANISM: Aspergillus fumigatus
US-10-128-714-3165

  Query Match           60.6%;  Score 40;  DB 15;  Length 794;
  Best Local Similarity  66.7%;  Pred. No. 49;
  Matches    8;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 QRFSKIASNTQS 13
              :|||||  |: ||
Db        639 ERFSKIVSSVQS 650


RESULT 4
US-10-128-714-8165
; Sequence 8165, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Bo
;  APPLICANT:  Hu, Wenqi
;  APPLICANT:  Tishkoff, Daniel
;  APPLICANT:  Zamudio, Carlos
;  APPLICANT:  Eroshkin, Alexey M
;  APPLICANT:  Lemieux, Sebastien M
;  TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus ar
;  TITLE OF INVENTION:  Methods of Use
;  FILE REFERENCE: 10182-018-999
;  CURRENT APPLICATION NUMBER: US/10/128,714
;  CURRENT FILING DATE:  2002-04-23
;  PRIOR APPLICATION NUMBER: US 60/285,697
;  PRIOR FILING DATE: 2001-04-23
;  PRIOR APPLICATION NUMBER: US 60/287,066
;  PRIOR FILING DATE: 2001-04-27
;  PRIOR APPLICATION NUMBER: US 60/295,890
;  PRIOR FILING DATE: 2001-06-05
;  PRIOR APPLICATION NUMBER: US 60/303,899
;  PRIOR FILING DATE: 2001-07-09
;  PRIOR APPLICATION NUMBER: US 60/316,362
;  PRIOR FILING DATE: 2001-08-31
;  NUMBER OF SEQ ID NOS: 8603
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8165
;   LENGTH: 1138
;   TYPE: PRT
;   ORGANISM: Aspergillus fumigatus
;   FEATURE:
;   NAME/KEY: MISC_FEATURE
;   LOCATION: (31)..(31)
;   OTHER INFORMATION: X= any amino acid
US-10-128-714-8165

  Query Match           60.6%;  Score 40;  DB 15;  Length 1138;
  Best Local Similarity  66.7%;  Pred. No. 72;
  Matches    8;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 QRFSKIASNTQS 13
              :|||||  |: ||
Db        951 ERFSKIVSSVQS 962


RESULT 5
US-10-106-698-5185
; Sequence 5185, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
;  APPLICANT: Ruben et al.
;  TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypep
;  FILE REFERENCE: PA005P1
```

```
;  CURRENT APPLICATION NUMBER: US/10/136,695
;  CURRENT FILING DATE:  2002-03-27
;  PRIOR APPLICATION NUMBER: PCT/US00/26524
;  PRIOR FILING DATE: 2000-09-28
;  PRIOR APPLICATION NUMBER: US 60/157,137
;  PRIOR FILING DATE: 1999-09-29
;  PRIOR APPLICATION NUMBER: US 60/163,280
;  PRIOR FILING DATE: 1999-11-03
;  NUMBER OF SEQ ID NOS: 8564
;  SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 5185
;   LENGTH: 78
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-106-698-5185

  Query Match            54.5%;  Score 36;  DB 15;  Length 78;
  Best Local Similarity  57.1%;  Pred. No. 22;
  Matches    8;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 SQRFSKIASNTQSR 14
              || |: :|  |||:
Db         21 SQSFNSMAQVTQSR 34


RESULT 5
US-09-923-844B-2
; Sequence 2, Application US/09923844B
; Patent No. US20020166143A1
; GENERAL INFORMATION:
;  APPLICANT: Pioneer Hi-Bred International, Inc.
;  APPLICANT:  Bao, Zhongmeng
;  APPLICANT:  Du, Guihua
;  TITLE OF INVENTION: Sclerotinia-inducible Genes and
;  TITLE OF INVENTION:  Promoters and Their Uses
;  FILE REFERENCE: 35718/234631
;  CURRENT APPLICATION NUMBER: US/09/923,844B
;  CURRENT FILING DATE:  2001-08-07
;  PRIOR APPLICATION NUMBER: US 60/224,603
;  PRIOR FILING DATE: 2000-08-11
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 371
;   TYPE: PRT
;   ORGANISM: Helianthus annuus
US-09-923-844B-2

  Query Match            53.0%;  Score 35;  DB 13;  Length 371;
  Best Local Similarity  61.5%;  Pred. No. 1.3e+02;
  Matches    8;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 QRFSKIASNTQSR 14
              | ||:|||:  ||
Db        105 QLFSKLASSPQSR 117


RESULT 7
US-09-881-752A-150
; Sequence 150, Application US/09881752A
; Patent No. US20020115078A1
; GENERAL INFORMATION:
;  APPLICANT: Kleanthous, Harold
;  APPLICANT:  Al-Garawi, Amal
;  APPLICANT:  Miller, Charles
;  APPLICANT:  Tomb, Jean-Francois
;  APPLICANT:  Oomen, Raymond P.
;  TITLE OF INVENTION: Identification of Polynucleotides
;  TITLE OF INVENTION:  Encoding No. US20020115078A1e1 Helicobacter Polypeptides in the
;  TITLE OF INVENTION:  Genome
;  FILE REFERENCE: 06132/041002
;  CURRENT APPLICATION NUMBER: US/09/881,752A
;  CURRENT FILING DATE:  2001-06-15
;  PRIOR APPLICATION NUMBER: US 08/833,457
;  PRIOR FILING DATE: 1997-04-01
;  NUMBER OF SEQ ID NOS: 370
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 150
;   LENGTH: 1230
;   TYPE: PRT
;   ORGANISM: Helicobacter pylori
US-09-881-752A-150

  Query Match            53.0%;  Score 35;  DB 10;  Length 1230;
  Best Local Similarity  61.5%;  Pred. No. 6.5e-02;
  Matches    8;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 SQRFSKIASNTQS 13
              ||  :|: |:|||
Db        136 SQITAKLLSNTQS 148


RESULT 8
US-09-896-923-40
; Sequence 40, Application US/09896923
; Publication No. US20030157483A1
; GENERAL INFORMATION:
;  APPLICANT: Sorge, Joseph
;  APPLICANT:  Hurlbut Hogrefe, Holly
;  APPLICANT:  Connie, Hansen
;  TITLE OF INVENTION: Compositions and Methods Utilizing DNA Polymerases
;  FILE REFERENCE: 25436/1560
;  CURRENT APPLICATION NUMBER: US/09/896,923
;  CURRENT FILING DATE:  2000-10-27
;  PRIOR APPLICATION NUMBER: 60/162,600
;  PRIOR FILING DATE: 1999-10-29
;  NUMBER OF SEQ ID NOS: 48
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 40
;   LENGTH: 180
;   TYPE: PRT
;   ORGANISM: Thermococcus sp. JDF-3
;   FEATURE:
;   NAME/KEY: Unsure
;   LOCATION: (114)..(114)
;   OTHER INFORMATION: X = Unknown
US-09-896-923-40

  Query Match            51.5%;  Score 34;  DB 12;  Length 180;
  Best Local Similarity  63.6%;  Pred. No. 1.2e+02;
  Matches    7;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 FSKIASNTQSR 14
              :|||:  |||:
Db        135 WSKIAKETQAR 145


RESULT 9
US-10-169-048-30
; Sequence 30, Application US/10169048
; Publication No. US20030072769A1
; GENERAL INFORMATION:
;  APPLICANT: Clarke, Edna Elizabeth
;  APPLICANT:  Zhou, Liqing
;  APPLICANT:  Shea, Jacqueline Elizabeth
;  APPLICANT:  Feldman, Robert Graham
;  APPLICANT:  Holden, David William
;  TITLE OF INVENTION: Streptococcus Pyogenes Virulence Genes and Proteins And Thei
;  FILE REFERENCE: GJE-97
;  CURRENT APPLICATION NUMBER: US/10/169,048
;  CURRENT FILING DATE:  2002-06-24
;  PRIOR APPLICATION NUMBER: PCT/GB00/04997
;  PRIOR FILING DATE: 2000-12-22
;  NUMBER OF SEQ ID NOS: 62
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 30
```

```
;   LENGTH: 407
;   TYPE: PRT
;   ORGANISM: Streptococcus pyogenes
US-10-159-048-30

  Query Match             51.5%;  Score 34;  DB 15;  Length 407;
  Best Local Similarity   58.3%;  Pred. No. 3e+02;
  Matches    7;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          3 RFSKIASNTQSR 14
              :|||: | |: :
Db        202 KFSKIVSLIEKR 213


RESULT 10
US-10-213-990-21
; Sequence 21, Application US/10213990
; Publication No. US20030082595A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Bo
;  APPLICANT:  Bussey, Howard
;  APPLICANT:  Storms, Reg
;  APPLICANT:  Roemer, Terry
;  TITLE OF INVENTION: NUCLEIC ACIDS OF ASPERGILLUS FUMIGATUS ENCODING INDUSTRIAL
;  TITLE OF INVENTION:  ENZYMES AND METHODS OF USE
;  FILE REFERENCE: 10182-019-999
;  CURRENT APPLICATION NUMBER: US/10/213,990
;  CURRENT FILING DATE:  2002-08-05
;  NUMBER OF SEQ ID NOS: 72
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 21
;   LENGTH: 704
;   TYPE: PRT
;   ORGANISM: Aspergillus
US-10-213-990-21

  Query Match             51.5%;  Score 34;  DB 15;  Length 704;
  Best Local Similarity   50.0%;  Pred. No. 5.4e+02;
  Matches    7;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 SQRFSKIASNTQSR 14
              :: |  |  :| ||:
Db        318 AQRLGKTCSNCQSQ 331


RESULT 11
US-09-801-368-98
; Sequence 98, Application US/09801368
; Patent No. US20020128250A1
; GENERAL INFORMATION:
;  APPLICANT: Busby, Robert
;  APPLICANT:  Cali, Brian
;  APPLICANT:  Hecht, Peter
;  APPLICANT:  Holtzman, Doug
;  APPLICANT:  Madden, Kevin
;  APPLICANT:  Maxon, Mary
;  APPLICANT:  Milne, Todd
;  APPLICANT:  No. US20020128250Alman, Thea
;  APPLICANT:  Royer, John
;  APPLICANT:  Salama, Sofie
;  APPLICANT:  Sherman, Amir
;  APPLICANT:  Silva, Jeff
;  APPLICANT:  Summers, Eric
;  TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
;  FILE REFERENCE: 109272.147
;  CURRENT APPLICATION NUMBER: US/09/801,368
;  CURRENT FILING DATE:  2001-03-07
;  PRIOR APPLICATION NUMBER: US 09/487,558
;  PRIOR FILING DATE: 2000-01-19
;  PRIOR APPLICATION NUMBER: US 60/160,587
;  PRIOR FILING DATE: 1999-10-20
;  NUMBER OF SEQ ID NOS: 440
;  SOFTWARE: Patentin version 3.0
; SEQ ID NO 98
;   LENGTH: 867
;   TYPE: PRT
;   ORGANISM: Aspergillus nidulans
US-09-801-368-98

  Query Match             51.5%;  Score 34;  DB 10;  Length 867;
  Best Local Similarity   66.7%;  Pred. No. 6.6e+02;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          6 KIASNTQSR 14
              :: :|||||
Db         11 KVGNNTQSR 19


RESULT 12
US-10-157-031-108
; Sequence 108, Application US/10157031
; Publication No. US20030108890A1
; GENERAL INFORMATION:
;  APPLICANT: Baranova, A. V.
;  APPLICANT:  Yankovsky, N. K.
;  APPLICANT:  Kozlov, A. P.
;  APPLICANT:  Lobashev, A. V.
;  APPLICANT:  Krukovskaya, L. L.
;  TITLE OF INVENTION: In silico screening for phenotype-associated expressed sequen
;  FILE REFERENCE: 2760-103
;  CURRENT APPLICATION NUMBER: US/10/157,031
;  CURRENT FILING DATE:  2002-05-30
;  NUMBER OF SEQ ID NOS: 415
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 108
;   LENGTH: 1042
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-157-031-108

  Query Match             51.5%;  Score 34;  DB 15;  Length 1042;
  Best Local Similarity   54.5%;  Pred. No. 8.3e+02;
  Matches    6;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 FSKIASNTQSR 14
              |:| |:||: :
Db        348 FSKFANNTEGQ 358


RESULT 13
US-10-156-761-13716
; Sequence 13716, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
;  APPLICANT: OMURA, SATOSHI
;  APPLICANT:  IKEDA, HARUO
;  APPLICANT:  ISHIKAWA, JUN
;  APPLICANT:  HORIKAWA, HIROSHI
;  APPLICANT:  SHIBA, TADAYOSHI
;  APPLICANT:  SAKAKI, YOSHIYUKI
;  APPLICANT:  HATTORI, MASAHIRA
;  TITLE OF INVENTION: NOVEL POLYNUCLECTIDES
;  FILE REFERENCE: 249-262
;  CURRENT APPLICATION NUMBER: US/10/156,761
;  CURRENT FILING DATE:  2002-05-29
;  PRIOR APPLICATION NUMBER: JP 2001-204089
;  PRIOR FILING DATE: 2001-05-30
;  PRIOR APPLICATION NUMBER: JP 2001-272697
;  PRIOR FILING DATE: 2001-08-02
;  NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 13716
;   LENGTH: 1515
;   TYPE: PRT
;   ORGANISM: Streptomyces avermitilis
US-10-156-761-13716
```

```
  Query Match             51.5%;  Score 34;  DB 15;  Length 1516;
  Best Local Similarity   63.6%;  Pred. No. 1.2e+03;
  Matches    7;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 SQRFSKIASNT 11
              :|:|.||  :|
Db       1474 AQRFSKIIPST 1484


RESULT 14
US-09-928-457-38
; Sequence 38, Application US/09928457
; Patent No. US20020154603A1
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  DNA, specific proteins and peptides
;    TITLE OF INVENTION:  of the Neisseria meningitidis species bacteria, method
;    TITLE OF INVENTION:  for obtaining them and their biological application.
;    NUMBER OF SEQUENCES:  98
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30 (OEB)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/928,457
;      FILING DATE:  2001-08-14
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/214,759
;      FILING DATE:  199-12-10
;  INFORMATION FOR SEQ ID NO:  38:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1981 acides amin,s
;      TYPE:  acide amin,
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  Peptide
;      LOCATION:  1..1981
US-09-928-457-38

  Query Match             51.5%;  Score 34;  DB 10;  Length 1981;
  Best Local Similarity   58.3%;  Pred. No. 1.7e+03;
  Matches    7;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          3 RFSKIASNTQSR 14
              |:|:|| :|  |
Db        808 RYSQIADNTTLR 819


RESULT 15
US-10-066-551-1
; Sequence 1, Application US/10066551
; Publication No. US20030100071A1
; GENERAL INFORMATION:
;  APPLICANT: Apicella, M. A.
;  APPLICANT:  Edwards, J. L.
;  APPLICANT:  Gibson, B. W.
;  APPLICANT:  Scheffler, K.
;  APPLICANT:  Brown, E.
;  TITLE OF INVENTION: Vaccine and compositions for the
;  TITLE OF INVENTION:  prevention and treatment of Neisserial infections
;  FILE REFERENCE: 875.045US1
;  CURRENT APPLICATION NUMBER: US/10/066,551
;  CURRENT FILING DATE:  2002-01-31
;  PRIOR APPLICATION NUMBER: US 60/344,452
;  PRIOR FILING DATE: 2001-10-23
;  PRIOR APPLICATION NUMBER: US 60/310,356
;  PRIOR FILING DATE: 2001-08-06
;  PRIOR APPLICATION NUMBER: US 60/266,070
;  PRIOR FILING DATE: 2001-01-31
;  NUMBER OF SEQ ID NOS: 12
; SEQ ID NO 1
;   LENGTH: 2015
;   TYPE: PRT
;   ORGANISM: Neisseria gonorrhoeae
US-10-066-551-1

  Query Match             51.5%;  Score 34;  DB 15;  Length 2015;
  Best Local Similarity   58.3%;  Pred. No. 1.7e+03;
  Matches    7;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          3 RFSKIASNTQSR 14
              |:|:|| :|  |
Db        808 RYSQIADNTTLR 819


Search completed: September 24, 2003, 14:22:31
Job time : 32.3034 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 24, 2003, 13:34:08 ; Search time 16.9858 Seconds
(without alignments)
79.250 Million cell updates/sec

Title: US-09-744-282-8
Perfect score: 66
Sequence: 1 SQRFSKIASNTQSR 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283306 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283306

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	66	100.0	383	2	S51327	heparin-binding gl
2	66	100.0	383	2	A49562	cartilage glycopro
3	46	69.7	405	2	S61551	breast-regressing
4	39	59.1	702	2	T13058	NADH2 dehydrogenas
5	39	59.1	703	2	T13393	NADH2 dehydrogenas
6	39	59.1	704	2	T13503	NADH2 dehydrogenas
7	39	59.1	706	2	T13391	NADH2 dehydrogenas
8	39	59.1	1244	2	C96584	hypothetical prote
9	38	57.6	84	2	T13498	NADH2 dehydrogenas
10	38	57.6	212	2	T13397	NADH2 dehydrogenas
11	38	57.6	212	2	T13723	NADH2 dehydrogenas
12	38	57.6	520	2	T06625	hypothetical prote
13	38	57.6	701	2	T13056	NADH2 dehydrogenas
14	38	57.6	702	2	T13409	NADH2 dehydrogenas
15	38	57.6	703	2	T12696	NADH2 dehydrogenas
16	38	57.6	706	2	T12748	NADH2 dehydrogenas
17	37	56.1	179	2	B72784	hypothetical prote
18	37	56.1	212	2	T13582	NADH2 dehydrogenas
19	37	56.1	577	2	B84582	hypothetical prote
20	37	56.1	702	2	T13505	NADH2 dehydrogenas
21	36	54.5	212	2	T13227	NADH2 dehydrogenas
22	36	54.5	474	1	A32386	protein-tyrosine-p
23	36	54.5	579	2	H88632	protein F56H3.7 [1
24	36	54.5	704	2	T13665	NADH2 dehydrogenas
25	36	54.5	886	2	S73184	phycobilisome link
26	36	54.5	918	2	B84580	hypothetical prote
27	36	54.5	1314	2	T09481	mating type silenc
28	35	53.0	212	2	T13571	NADH2 dehydrogenas
29	35	53.0	311	1	JC1191	osteopontin precur
30	35	53.0	457	2	T13109	hypothetical prote
31	35	53.0	469	2	T40339	ferredoxin-NADP re
32	35	53.0	495	2	T04784	hypothetical prote
33	35	53.0	524	2	T08931	hypothetical prote
34	35	53.0	582	2	D82146	transport ATP-bind
35	35	53.0	633	2	S45177	BEM4 protein - yea
36	35	53.0	660	2	A71423	hypothetical prote
37	35	53.0	701	2	T13587	NADH2 dehydrogenas
38	35	53.0	792	2	T49989	hypothetical prote
39	35	53.0	808	2	S73947	preprotein translo
40	35	53.0	1230	2	B64664	outer membrane pro
41	35	53.0	1237	2	D71850	probable outer mem
42	34.5	52.3	395	2	S43746	MNN9 protein - yea
43	34	51.5	59	2	T12647	NADH2 dehydrogenas
44	34	51.5	159	2	T10826	probable resistanc
45	34	51.5	207	2	T12662	NADH2 dehydrogenas

ALIGNMENTS

RESULT 1
S51327
heparin-binding glycoprotein 38K - pig
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 22-Jun-1999
C;Accession: S51327
R;Shackelton, L.M.; Mann, D.M.; Millis, A.J.T.
submitted to the EMBL Data Library, January 1995
A;Description: Identification of a 38kDa heparin-binding glycoprotein (gp38k) in di
A;Reference number: S51327
A;Accession: S51327
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-383 <SHA>
A;Cross-references: EMBL:Z47803; NID:g634097; PIDN:CAA87764.1; PID:g634098
C;Superfamily: Streptomyces chitinase chi40

Query Match 100.0%; Score 66; DB 2; Length 383;
Best Local Similarity 100.0%; Pred. No. 4.2e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SQRFSKIASNTQSR 14
||||||||||||||
Db 103 SQRFSKIASNTQSR 116

RESULT 2
A49562
cartilage glycoprotein gp39 precursor - human
N;Alternate names: 39K synovial protein
C;Species: Homo sapiens (man)
C;Date: 23-Mar-1995 #sequence_revision 23-Mar-1995 #text_change 22-Jun-1999
C;Accession: A49562; S10677; A33182
R;Hakala, B.E.; White, C.; Recklies, A.D.
J. Biol. Chem. 268, 25803-25810, 1993
A;Title: Human cartilage gp-39, a major secretory product of articular chondrocytes
A;Reference number: A49562; MUID:94064659; PMID:8245017
A;Accession: A49562
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-383 <HAK>
A;Cross-references: GB:M80927; NID:g348911; PIDN:AAA16074.1; PID:g348912
R;Nyirkos, P.; Golds, E.E.
Biochem. J. 269, 265-268, 1990
A;Title: Human synovial cells secrete a 39 kDa protein similar to a bovine mammary
A;Reference number: S10677; MUID:90328983; PMID:2375755
A;Accession: S10677
A;Molecule type: protein
A;Residues: 22-40,'X',42-45 <NY2>
C;Superfamily: Streptomyces chitinase chi40
C;Keywords: cartilage; extracellular protein; glycoprotein
F;1-21/Domain: signal sequence #status predicted <SIG>

```
F;22-383/Product: cartilage glycoprotein gp39 #status predicted <MAT>

  Query Match           100.0%;  Score 66;  DB 2;  Length 383;
  Best Local Similarity  100.0%;  Pred. No. 4.2e-05;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SQRFSKIASNTQSR 14
              ||||||||||||||
Db        103 SQRFSKIASNTQSR 116


RESULT 3
S61551
breast-regressing protein brp39 precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
C;Accession: S61551; S61550; I48271
R;Morrison, B.W.; Leder, P.
Oncogene 9, 3417-3426, 1994
A;Title: neu and ras initiate murine mammary tumors that share genetic markers generally
A;Reference number: I48271; MUID:95060797; PMID:7970709
A;Accession: S61551
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-405 <MOR1>
A;Cross-references: EMBL:X93035; NID:g1085065; PIDN:CAA63603.1; PID:g1085066
R;Morrison, B.W.
submitted to the EMBL Data Library, November 1995
A;Reference number: S61550
A;Accession: S61550
A;Molecule type: mRNA
A;Residues: 1-245,'T',247-330,'H',332-350,'MVWALDLDDFQGTCQPKEFFPLTNAIKDALA' <MOR2>
A;Cross-references: EMBL:X93035; NID:g1085065; PIDN:CAA63603.1; PID:g1085066
A;Note: the differences at the carboxyl end are due to a frameshift error
C;Genetics:
A;Gene: brp39
C;Superfamily: Streptomyces chitinase ch140
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-405/Product: breast-regressing protein brp39 #status predicted <MAT>

  Query Match           69.7%;  Score 46;  DB 2;  Length 405;
  Best Local Similarity  69.2%;  Pred. No. 0.47;
  Matches    9;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 QRFSKIASNTQSR 14
              :|||:|||||| :
Db        105 KRFSEIASNTERR 117


RESULT 4
T13058
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 5 - perennial veldt grass chloroplas
C;Species: chloroplast Ehrharta calycina (perennial veldt grass)
C;Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 03-Jun-2002
C;Accession: T13058
R;Clark, L.G.; Zhang, W.; Wendel, J.F.
Syst. Bot. 20, 436-460, 1995
A;Title: A phylogeny of the grass family (Poaceae) based on ndhF-sequence data.
A;Reference number: Z17570
A;Accession: T13058
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-702 <CLA>
A;Cross-references: EMBL:U21995; NID:g726155; PID:g726156; PIDN:AAA64209.1
A;Experimental source: leaf
C;Genetics:
A;Genome: chloroplast
A;Note: ndhF
C;Superfamily: NADH dehydrogenase (ubiquinone) chain 5
C;Keywords: chloroplast; membrane-associated complex; NAD; oxidoreductase

  Query Match           59.1%;  Score 39;  DB 2;  Length 702;
  Best Local Similarity  57.1%;  Pred. No. 18;
  Matches    8;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 SQRFSKIASNTQSR 14
              |:  ||| |:|::|
Db        496 SKNISKIPSNTRNR 509


RESULT 5
T13393
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 5 - Lithachne pauciflora chlorop
C;Species: chloroplast Lithachne pauciflora
C;Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 03-Jun-2002
C;Accession: T13393
R;Clark, L.G.; Zhang, W.; Wendel, J.F.
Syst. Bot. 20, 436-460, 1995
A;Title: A phylogeny of the grass family (Poaceae) based on ndhF-sequence data.
A;Reference number: Z17570
A;Accession: T13393
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-703 <CLA>
A;Cross-references: EMBL:U21978; NID:g755846; PID:g755847; PIDN:AAA64697.1
A;Experimental source: leaf
C;Genetics:
A;Genome: chloroplast
A;Note: ndhF
C;Superfamily: NADH dehydrogenase (ubiquinone) chain 5
C;Keywords: chloroplast; membrane-associated complex; NAD; oxidoreductase

  Query Match           59.1%;  Score 39;  DB 2;  Length 703;
  Best Local Similarity  57.1%;  Pred. No. 18;
  Matches    8;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 SQRFSKIASNTQSR 14
              |:  |||  |:::|
Db        495 SQNISKIKGNTRNR 508


RESULT 6
T13503
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 5 - Olyra latifolia chloroplast
C;Species: chloroplast Olyra latifolia
C;Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 03-Jun-2002
C;Accession: T13503
R;Clark, L.G.; Zhang, W.; Wendel, J.F.
Syst. Bot. 20, 436-460, 1995
A;Title: A phylogeny of the grass family (Poaceae) based on ndhF-sequence data.
A;Reference number: Z17570
A;Accession: T13503
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-704 <CLA>
A;Cross-references: EMBL:U21971; NID:g755834; PID:g755835; PIDN:AAA64776.1
A;Experimental source: leaf
C;Genetics:
A;Genome: chloroplast
A;Note: ndhF
C;Superfamily: NADH dehydrogenase (ubiquinone) chain 5
C;Keywords: chloroplast; membrane-associated complex; NAD; oxidoreductase

  Query Match           59.1%;  Score 39;  DB 2;  Length 704;
  Best Local Similarity  57.1%;  Pred. No. 18;
  Matches    8;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 SQRFSKIASNTQSR 14
              |:  |||  |:::|
Db        496 SQNISKIQGNTRNR 509


RESULT 7
T13391
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 5 - Lithachne humilis chloroplas
C;Species: chloroplast Lithachne humilis
```

```
C;Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 03-Jun-2002
C;Accession: T13391
R;Clark, L.G.; Zhang, W.; Wendel, J.F.
Syst. Bot. 20, 436-460, 1995
A;Title: A phylogeny of the grass family (Poaceae) based on ndhF-sequence data.
A;Reference number: Z17570
A;Accession: T13391
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-706 <CLA>
A;Cross-references: EMBL:U21977; NID:g755844; PID:g755845; PIDN:AAA64896.1
A;Experimental source: leaf
C;Genetics:
A;Genome: chloroplast
A;Note: ndhF
C;Superfamily: NADH dehydrogenase (ubiquinone) chain 5
C;Keywords: chloroplast; membrane-associated complex; NAD; oxidoreductase

  Query Match             59.1%;  Score 39;  DB 2;  Length 706;
  Best Local Similarity   57.1%;  Pred. No. 18;
  Matches    8;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 SQRFSKIASNTQSR 14
              || |||  ||:::
Db        495 SQNISKIKGNTRNR 508


RESULT 8
C96584
hypothetical protein F20D21.10 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C;Accession: C96584
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Ching, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: C96584
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1244 <STO>
A;Cross-references: GB:AC005173; NID:g6554464; PIDN:AAD25608.2; GSPDB:GN00141
C;Genetics:
A;Gene: F20D21.10
A;Map position: 1

  Query Match             59.1%;  Score 39;  DB 2;  Length 1244;
  Best Local Similarity   53.8%;  Pred. No. 34;
  Matches    7;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 SQRFSKIASNTQS 13
              :|||:|:|| | :
Db        484 TQRYAKLASKTSS 496


RESULT 9
T13498
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 5 - Nardus stricta chloroplast (frag
C;Species: chloroplast Nardus stricta
C;Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 03-Jun-2002
C;Accession: T13498
R;Catalan, P.; Kellogg, E.A.; Olmstead, R.G.
Mol. Phylogenet. Evol. 8, 150-166, 1997
A;Title: Phylogeny of Poaceae subfamily Pooideae based on chloroplast ndhF gene sequenci
A;Reference number: Z15553; MUID:97446273; PMID:9299221
A;Accession: T13498
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-84 <CAT>
A;Cross-references: EMBL:U71051; NID:g1763168; PID:g1763169; PIDN:AAB39681.1
C;Genetics:
A;Genome: chloroplast
A;Note: ndhF
C;Superfamily: NADH dehydrogenase (ubiquinone) chain 5
C;Keywords: chloroplast; membrane-associated complex; NAD; oxidoreductase

  Query Match             57.6%;  Score 38;  DB 2;  Length 84;
  Best Local Similarity   57.1%;  Pred. No. 2.9;
  Matches    8;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 SQRFSKIASNTQSR 14
              || |||  ||:::
Db         44 SQNISKIHGNTRNR 57


RESULT 10
T13397
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 5 - Lygeum spartum chloroplast (
C;Species: chloroplast Lygeum spartum
C;Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 03-Jun-2002
C;Accession: T13397
R;Catalan, P.; Kellogg, E.A.; Olmstead, R.G.
Mol. Phylogenet. Evol. 8, 150-166, 1997
A;Title: Phylogeny of Poaceae subfamily Pooideae based on chloroplast ndhF gene seq
A;Reference number: Z15553; MUID:97446273; PMID:9299221
A;Accession: T13397
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-212 <CAT>
A;Cross-references: EMBL:U71050; NID:g1763164; PID:g1763165; PIDN:AAB39679.1
C;Genetics:
A;Genome: chloroplast
A;Note: ndhF
C;Superfamily: NADH dehydrogenase (ubiquinone) chain 5
C;Keywords: chloroplast; membrane-associated complex; NAD; oxidoreductase

  Query Match             57.6%;  Score 38;  DB 2;  Length 212;
  Best Local Similarity   57.1%;  Pred. No. 7.8;
  Matches    8;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 SQRFSKIASNTQSR 14
              || |||  ||:::
Db         44 SQNISKIHGNTRNR 57


RESULT 11
T13723
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 5 - Stipa parviflora chloroplast
C;Species: chloroplast Stipa parviflora
C;Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 03-Jun-2002
C;Accession: T13723
R;Catalan, P.; Kellogg, E.A.; Olmstead, R.G.
Mol. Phylogenet. Evol. 8, 150-166, 1997
A;Title: Phylogeny of Poaceae subfamily Pooideae based on chloroplast ndhF gene seq
A;Reference number: Z15553; MUID:97446273; PMID:9299221
A;Accession: T13723
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-212 <CAT>
A;Cross-references: EMBL:U71045; NID:g1763186; PID:g1763187; PIDN:AAB39690.1
C;Genetics:
A;Genome: chloroplast
A;Note: ndhF
C;Superfamily: NADH dehydrogenase (ubiquinone) chain 5
C;Keywords: chloroplast; membrane-associated complex; NAD; oxidoreductase

  Query Match             57.6%;  Score 38;  DB 2;  Length 212;
  Best Local Similarity   57.1%;  Pred. No. 7.8;
  Matches    8;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;
```

```
Qy          1 SQRFSKIASNTQSR 14
              || |:|  |:::|
Db         44 SQNLSKIHGNTRNR 57


RESULT 12
T06625
hypothetical protein T20K18.20 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 22-Oct-1999
C;Accession: T06625
R;Bevan, M.; Peters, S.A.; van Staveren, M.; Dirkse, W.; Stiekema, W.; Bancroft, I.; Mew
submitted to the Protein Sequence Database, April 1999
A;Reference number: Z15790
A;Accession: T06625
A;Molecule type: DNA
A;Residues: 1-520 <BEV>
A;Cross-references: EMBL:AL049640; GSPDB:GN00062; ATSP:T20K18.20
A;Experimental source: cultivar Columbia; BAC clone T20K18
C;Genetics:
A;Gene: ATSP:T20K18.20
A;Map position: 4
A;Introns: 418/3; 450/3

  Query Match             57.6%;  Score 38;  DB 2;  Length 520;
  Best Local Similarity   50.0%;  Pred. No. 21;
  Matches    7;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 SQRFSKIASNTQSR 14
              |:|  |:::| | |:
Db         65 SQKMQKVAAETHSR 78


RESULT 13
T13056
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 5 - weeping love grass chloroplast (
C;Species: chloroplast Eragrostis curvula (weeping love grass)
C;Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 03-Jun-2002
C;Accession: T13056
R;Clark, L.G.; Zhang, W.; Wendel, J.F.
Syst. Bot. 20, 436-460, 1995
A;Title: A phylogeny of the grass family (Poaceae) based on ndhF-sequence data.
A;Reference number: Z17570
A;Accession: T13056
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-701 <CLA>
A;Cross-references: EMBL:U21989; NID:g726141; PID:g726142; PIDN:AAA64193.1
A;Experimental source: leaf
C;Genetics:
A;Genome: chloroplast
A;Note: ndhF
C;Superfamily: NADH dehydrogenase (ubiquinone) chain 5
C;Keywords: chloroplast; membrane-associated complex; NAD; oxidoreductase

  Query Match             57.6%;  Score 38;  DB 2;  Length 701;
  Best Local Similarity   57.1%;  Pred. No. 29;
  Matches    8;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 SQRFSKIASNTQSR 14
              :|   :| |||:::|
Db        496 SQNIPKIPSNTRNR 509


RESULT 14
T13409
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 5 - Micraira lazaridisii chloroplast
C;Species: chloroplast Micraira lazaridisii
C;Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 03-Jun-2002
C;Accession: T13409
R;Clark, L.G.; Zhang, W.; Wendel, J.F.
Syst. Bot. 20, 436-460, 1995
A;Title: A phylogeny of the grass family (Poaceae) based on ndhF-sequence data.
A;Reference number: Z17570
A;Accession: T13409
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-702 <CLA>
A;Cross-references: EMBL:U21972; NID:g755836; PID:g755837; PIDN:AAA64777.1
A;Experimental source: leaf
C;Genetics:
A;Genome: chloroplast
A;Note: ndhF
C;Superfamily: NADH dehydrogenase (ubiquinone) chain 5
C;Keywords: chloroplast; membrane-associated complex; NAD; oxidoreductase

  Query Match             57.6%;  Score 38;  DB 2;  Length 702;
  Best Local Similarity   57.1%;  Pred. No. 29;
  Matches    8;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 SQRFSKIASNTQSR 14
              |:  |||  |:::|
Db        497 SQNISKIHGNTRNR 510


RESULT 15
T12696
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 5 - Brachyelytrum erectum chloro
C;Species: chloroplast Brachyelytrum erectum
C;Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 03-Jun-2002
C;Accession: T12696
R;Clark, L.G.; Zhang, W.; Wendel, J.F.
submitted to the EMBL Data Library, February 1995
A;Description: A phylogeny of the Grass family based on ndhF sequence data.
A;Reference number: Z14220
A;Accession: T12696
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-703 <CLA>
A;Cross-references: EMBL:U22005; NID:g726173; PID:g726174; PIDN:AAA64209.1
A;Experimental source: leaf
C;Genetics:
A;Genome: chloroplast
A;Note: ndhF
C;Superfamily: NADH dehydrogenase (ubiquinone) chain 5
C;Keywords: chloroplast; membrane-associated complex; NAD; oxidoreductase

  Query Match             57.6%;  Score 38;  DB 2;  Length 703;
  Best Local Similarity   57.1%;  Pred. No. 29;
  Matches    8;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 SQRFSKIASNTQSR 14
              |:  |||  |:::|
Db        496 SQNISKIHGNTRNR 509


Search completed: September 24, 2003, 13:49:26
Job time : 17.9888 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 24, 2003, 13:26:37 ; Search time 8.80899 Seconds
(without alignments)
74.739 Million cell updates/sec

Title: US 09-744-282-8
Perfect score: 66
Sequence: 1 SQRFSKIASNTQSR 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	66	100.0	383	1	C3L1_HUMAN	P36222 homo sapien
2	46	69.7	381	1	C3L1_MOUSE	Q61362 mus musculu
3	39	59.1	1244	1	ALA6_ARATH	O9slk6 arabidopsis
4	36	54.5	259	1	MOBD_BPI4	P39241 bacteriopha
5	36	54.5	400	1	HFLK_VIBPA	P40605 vibrio para
6	36	54.5	554	1	MPIP_YEAST	P23748 saccharomyc
7	36	54.5	886	1	APCE_PORPU	P51263 porphyra pu
8	35	53.0	311	1	OSTP_RABIT	P31097 oryctolagus
9	35	53.0	633	1	BEM4_YEAST	P39021 saccharomyc
10	35	53.0	808	1	SECA_MYCPN	P75559 mycoplasma
11	34.5	52.3	394	1	MNN9_YEAST	P39107 saccharomyc
12	34	51.5	439	1	SLAP_LACHE	P38059 lactobacill
13	34	51.5	454	1	AAIM_LUPAN	P26563 lupinus ang
14	34	51.5	467	1	FLII_BUCAI	P57178 buchnera ap
15	34	51.5	705	1	NU5C_HORVU	Q32440 hordeum vu.
16	34	51.5	733	1	NU5C_WHEAT	Q95h46 triticum ae
17	34	51.5	806	1	SECA_MYCGE	P47318 mycoplasma
18	34	51.5	893	1	KNS8_YEAST	P53739 saccharomyc
19	34	51.5	1233	1	MU5A_HUMAN	P98088 homo sapien
20	33	50.0	201	1	DRPF_CRAPL	P25283 craterostig
21	33	50.0	296	1	COPE_YEAST	P40509 saccharomyc
22	33	50.0	340	1	VC01_SPVKA	P32281 swinepox vi
23	33	50.0	366	1	T2A1_ACICA	P24546 acinetobact
24	33	50.0	370	1	SERC_RABIT	P10658 oryctolagus
25	33	50.0	472	1	GSA2_ARATH	Q42522 arabidopsis
26	33	50.0	626	1	HTPG_BUCBP	Q89a93 buchnera ap
27	33	50.0	628	1	NXF1_CAEEL	Q9xvs7 caenorhabdi
28	33	50.0	727	1	PEC1_MOUSE	Q08481 mus musculu
29	33	50.0	734	1	NU5C_ORYSA	P12129 oryza sativ
30	33	50.0	738	1	PEC1_HUMAN	P16284 homo sapien
31	33	50.0	746	1	FEPA_ECOLI	P05825 escherichia
32	33	50.0	859	1	SYL_BUCAI	P57519 buchnera ap
33	33	50.0	1184	1	FBL2_HUMAN	P98095 homo sapien
34	33	50.0	1398	1	TOP2_PLAFK	P41001 plasmodium
35	32	48.5	208	1	UPP_VIBPA	Q87mhi vibrio para
36	32	48.5	272	1	RSR1_YEAST	P13856 saccharomyc
37	32	48.5	273	1	TYRO_COTJA	Q08410 coturnix co
38	32	48.5	286	1	PPG_CORGL	Q8nnv7 corynebacte
39	32	48.5	293	1	RRPP_CHAV	P16380 chandipura
40	32	48.5	346	1	UFE1_YEAST	P41834 saccharomyc
41	32	48.5	370	1	SERC_HUMAN	Q9y617 homo sapien
42	32	48.5	381	1	Y926_HELPY	P55985 helicobacte
43	32	48.5	416	1	HMT1_DROME	P22711 drosophila
44	32	48.5	474	1	SY65_DROME	P21521 drosophila
45	32	48.5	493	1	GLO3_YEAST	P38662 saccharomyc

ALIGNMENTS

RESULT 1
C3L1_HUMAN
ID C3L1_HUMAN STANDARD; PRT; 383 AA.
AC P36222; P30923;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Chitinase-3 like protein 1 precursor (Cartilage glycoprotein-39) (GP-
DE 39) (39 kDa synovial protein) (YKL-40).
GN CHI3L1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Cartilage;
RX MEDLINE=94064658; PubMed=8245017;
RA Hakala B.E., White C., Recklies A.D.;
RT "Human cartilage gp-39, a major secretory product of articular
RT chondrocytes and synovial cells, is a mammalian member of a chitinase
RT protein family.";
RL J. Biol. Chem. 268:25803-25810(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=97386591; PubMed=9244440;
RA Rehli M., Krause S.W., Andreesen R.;
RT "Molecular characterization of the gene for human cartilage gp-39
RT (CHI3L1), a member of the chitinase protein family and marker for
RT late stages of macrophage differentiation.";
RL Genomics 43:221-225(1997).
RN [3]
RP SEQUENCE OF 22-45.
RX MEDLINE=90328983; PubMed=2375755;
RA Nyirkos P., Golds E.E.;
RT "Human synovial cells secrete a 39 kDa protein similar to a bovine
RT mammary protein expressed during the non-lactating period.";
RL Biochem. J. 269:265-268(1990).
CC -!- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN THE CAPACITY OF CELLS TO
CC RESPOND TO AND COPE WITH CHANGES IN THEIR ENVIRONMENT.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- TISSUE SPECIFICITY: PRESENT IN ARTICULAR CHONDROCYTES, SYNOVIAL
CC CELLS AS WELL AS IN LIVER. UNDETECTABLE IN MUSCLE TISSUES, LUNG,
CC PANCREAS, MONONUCLEAR CELLS, OR FIBROBLASTS.
CC -!- PTM: GLYCOSYLATED.
CC -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.
CC ----------

CC ----------

```
DR   EMBL; M80927; AAA16074.1; -.
DR   EMBL; Y08374; CAA69661.1; -.
DR   EMBL; Y08375; CAA69661.1; JOINED.
DR   EMBL; Y08376; CAA69661.1; JOINED.
DR   EMBL; Y08377; CAA69661.1; JOINED.
DR   EMBL; Y08378; CAA69661.1; JOINED.
DR   PIR; A49562; A49562.
DR   PDB; 1LA7; 10-APR-02.
DR   Genew; HGNC:1932; CHI3L1.
DR   MIM; 601525; -.
DR   GO; GO:0005578; C:extracellular matrix; TAS.
DR   GO; GO:0005615; C:extracellular space; TAS.
DR   GO; GO:0005207; F:extracellular matrix glycoprotein; TAS.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   InterPro; IPR001579; Glyco_hydro_18/2.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   ProDom; PD000471; Glyco_hydro_18; 1.
DR   SMART; SM00636; Glyco_18; 1.
DR   PROSITE; PS01095; CHITINASE_18; FALSE_NEG.
KW   Glycoprotein; Signal; 3D-structure.
FT   SIGNAL        1     21
FT   CHAIN        22    383       CHITINASE-3 LIKE PROTEIN 1.
FT   CARBOHYD     60     60       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   383 AA;  42613 MW;  76ADD8298EEEC2D1 CRC64;

  Query Match             100.0%;  Score 66;  DB 1;  Length 383;
  Best Local Similarity   100.0%;  Pred. No. 2.3e-05;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SQRFSKIASNTQSR 14
              ||:|||||||||||
Db        103 SQRFSKIASNTQSR 116


RESULT 2
C3L1_MOUSE
ID   C3L1_MOUSE     STANDARD;      PRT;   381 AA.
AC   Q61362;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Chitinase-3 like protein 1 precursor (Cartilage glycoprotein-39)
DE   (GP-39) (BRP39 protein).
GN   CHI3L1 OR BRP39.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FVB/N; TISSUE=Breast;
RX   MEDLINE=95060797; PubMed=7970700;
RA   Morrison B.W., Leder P.;
RT   "neu and ras initiate murine mammary tumors that share genetic markers
RT   generally absent in c-myc and int-2-initiated tumors.";
RL   Oncogene 9:3417-3426(1994).
CC   -!- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN THE CAPACITY OF CELLS TO
CC       RESPOND TO AND COPE WITH CHANGES IN THEIR ENVIRONMENT (BY
CC       SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Extracellular (By similarity).
CC   -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X93035; CAA63603.1; -.
DR   PIR; S61551; S61551.
DR   MGD; MGI:1340899; Chi3l1.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   InterPro; IPR001579; Glyco_hydro_18/2.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   ProDom; PD000471; Glyco_hydro_18; 1.
DR   SMART; SM00636; Glyco_18; 1.
DR   PROSITE; PS01095; CHITINASE_18; FALSE_NEG.
KW   Glycoprotein; Signal.
FT   SIGNAL        1     21       POTENTIAL.
FT   CHAIN        22    381       CHITINASE-3 LIKE PROTEIN 1.
FT   CARBOHYD     60     60       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   381 AA;  43001 MW;  EF6581E8184F0450 CRC64;

  Query Match             69.7%;  Score 46;  DB 1;  Length 381;
  Best Local Similarity   69.2%;  Pred. No. 0.19;
  Matches    9;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 QRFSKIASNTQSR 14
              :|||:|||||: |
Db        105 KRFSEIASNTERR 117


RESULT 3
ALA6_ARATH
ID   ALA6_ARATH     STANDARD;      PRT;  1244 AA.
AC   Q9SLK6;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Potential phospholipid-transporting ATPase 6 (EC 3.6.3.1)
DE   (Aminophospholipid flippase 6).
GN   ALA6 OR AT1G54280 OR F20D21.10.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RX   MEDLINE=21016719; PubMed=11130712;
RA   Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA   White C., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
RA   Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA   Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA   Dunn P., Etgu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA   Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA   Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA   Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA   Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA   Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziali A.,
RA   Militscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA   Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA   Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA   Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA   Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA   Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT   "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT   thaliana.";
RL   Nature 408:816-820(2000).
CC   -!- FUNCTION: INVOLVED IN TRANSPORT OF PHOSPHOLIPIDS (POTENTIAL).
CC   -!- CATALYTIC ACTIVITY: ATP + H(2)O = ADP + phosphate.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: Belongs to the cation transport ATPases family (P-type
CC       ATPases). Subfamily IV.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
```

```
DR   EMBL; AC005287; AAD25608.2; -.
DR   PIR; C96584; C96584.
DR   InterPro; IPR001757; ATPase_E1-E2.
DR   InterPro; IPR006539; Flippase.
DR   InterPro; IPR005834; Hydrolase.
DR   Pfam; PF00702; Hydrolase; 1.
DR   PRINTS; PR00119; CATATPASE.
DR   TIGRFAMs; TIGR01652; ATPase-Plipid; 1.
DR   TIGRFAMs; TIGR01494; ATPase_P-type; 7.
DR   PROSITE; PS00154; ATPASE_E1_E2; 1.
KW   Hydrolase; Transmembrane; Phosphorylation; ATP-binding; Metal-binding;
KW   Magnesium; Multigene family.
FT   DOMAIN        1     75       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     76     97       POTENTIAL.
FT   DOMAIN       98    101       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    102    124       POTENTIAL.
FT   DOMAIN      125    326       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    327    328       POTENTIAL.
FT   DOMAIN      329    360       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    361    378       POTENTIAL.
FT   DOMAIN      379    947       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    948    967       POTENTIAL.
FT   DOMAIN      968    981       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    982   1001       POTENTIAL.
FT   DOMAIN     1002   1031       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   1032   1054       POTENTIAL.
FT   DOMAIN     1055   1067       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   1068   1090       POTENTIAL.
FT   DOMAIN     1091   1096       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   1097   1117       POTENTIAL.
FT   DOMAIN     1118   1134       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   1135   1159       POTENTIAL.
FT   DOMAIN     1160   1222       CYTOPLASMIC (POTENTIAL).
FT   MOD_RES     426    426       PHOSPHORYLATION (BY SIMILARITY).
FT   METAL       892    892       MAGNESIUM (BY SIMILARITY).
FT   METAL       896    896       MAGNESIUM (BY SIMILARITY).
SQ   SEQUENCE   1244 AA;  140681 MW;  A3DC0954F965CDFE CRC64;

  Query Match           59.1%;  Score 39;  DB 1;  Length 1244;
  Best Local Similarity  53.8%;  Pred. No. 16;
  Matches    7;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 SQRFSKIASNTQS 13
              :|:::|:|: | |
Db        484 TQRYAKLASKTSS 496


RESULT 4
MOBD_BPT4
ID   MOBD_BPT4      STANDARD;      PRT;   259 AA.
AC   P39241; Q96216;
DT   01-FEB-1995 (Rel. 31, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Probable mobile endonuclease D.
GN   MOBD OR TK.-10.
OS   Bacteriophage T4.
OC   Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;
OC   T4-like viruses.
OX   NCBI_TaxID=10665;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Mzhavia N., Marusich E., Djavakhishvili T., Keitzel C., Peterson S.,
RA   Awaya M., Eidermiller J., Canada D., Tracy J., Gailbreath K.,
RA   Paddison P., Anderson B., Stidham T., Blattner F., Kutter E.M.;
RT   "The 10.7 kb 'nonessential' region of bacteriophage T4 between the
RT   genes tk and nrdC: twenty new T4 genes, generally conserved among
RT   T-even phages.";
RL   Submitted (NOV-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Kutter E., Arisaka F., Kunisawa T., Tsugita A., Mosig G.,
RA   Mesyanzhinov V., Ruger W., Stidham T., Thomas E.;
RT   "Bacteriophage T4 genome analysis.";
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: TO PHAGE T4 MOBB AND MOBC.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U76612; AAB26970.1; -.
DR   EMBL; AF158101; AAD42592.1; -.
DR   InterPro; IPR002711; HNH.
DR   InterPro; IPR003615; HNH_nuc.
DR   Pfam; PF01844; HNH; 1.
DR   SMART; SM00507; HNHc; 1.
KW   Hydrolase; Nuclease; Endonuclease.
SQ   SEQUENCE   259 AA;  30453 MW;  C149B7657EA138F4 CRC64;

  Query Match           54.5%;  Score 36;  DB 1;  Length 259;
  Best Local Similarity  70.0%;  Pred. No. 11;
  Matches    7;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 QRFSKIASNT 11
              ::|:|| ||
Db        103 EKFSKIHSNT 112


RESULT 5
HFLK_VIBPA
ID   HFLK_VIBPA     STANDARD;      PRT;   400 AA.
AC   P40605;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   HflK protein.
GN   HFLK OR VP2815.
OS   Vibrio parahaemolyticus.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC   Vibrionaceae; Vibrio.
OX   NCBI_TaxID=670;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BB22;
RA   McCarter L.L.;
RL   Submitted (APR-1994) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=RIMD 2210633 / Serotype O3:K6;
RX   MEDLINE=22508454; PubMed=12620739;
RA   Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
RA   Iijima Y., Najima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,
RA   Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;
RT   "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
RT   distinct from that of V. cholerae.";
RL   Lancet 361:743-749(2003).
CC   -!- FUNCTION: HFLC AND HFLK COULD ENCODE OR REGULATE A PROTEASE (BY
CC       SIMILARITY).
CC   -!- SUBUNIT: HFLC AND HFLK MAY INTERACT TO FORM A MULTIMERIC COMPLEX.
CC   -!- SIMILARITY: BELONGS TO THE BAND 7 / MEC-2 FAMILY. HFLK SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U09005; AAA62186.1; -.
DR   EMBL; AP005082; BAC61078.1; -.
```

```
DR   InterPro; IPR001107; Band_7.
DR   InterPro; IPR001972; Stomatin.
DR   Pfam; PF01145; Band_7; 1.
DR   PRINTS; PR00721; STOMATIN.
DR   SMART; SM00244; PHB; 1.
KW   Transmembrane; Complete proteome.
FT   TRANSMEM     70     90       POTENTIAL.
SQ   SEQUENCE   400 AA;  44237 MW;  B94D8363FD81DC94 CRC64;

  Query Match             54.5%;  Score 36;  DB 1;  Length 400;
  Best Local Similarity   57.1%;  Pred. No. 18;
  Matches    8;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 SQRFSKIASNTQSR 14
              :|  :  ||||:'
Db        378 SERTQRETSNTQSR 391


RESULT 6
MPIP_YEAST
ID   MPIP_YEAST     STANDARD;      PRT;   554 AA.
AC   P23748;
DT   01-NOV-1991 (Rel. 20, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   M-phase inducer phosphatase (EC 3.1.3.48) (Mitosis initiation protein
DE   MIH1) (Mitotic inducer homolog).
GN   MIH1 OR YMR036C OR YM9532.01C OR YM9973.10C.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89195240; PubMed=2649252;
RA   Russell P., Moreno S., Reed S.I.;
RT   "Conservation of mitotic controls in fission and budding yeasts.";
RL   Cell 57:295-303(1989).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288c / AB972;
RX   PubMed=9169872;
RA   Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,
RA   Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,
RA   Jagels K., Lye G., Moule S., Odell C., Pearson D., Rajandream M.A.,
RA   Rice P., Skelton J., Walsh S., Whitehead S., Barrell B.G.;
RT   "The nucleotide sequence of Saccharomyces cerevisiae chromosome
RT   XIII.";
RL   Nature 387:90-93(1997).
CC   -!- FUNCTION: Terminates the cell cycle delay. Reverses the CDC28
CC       phosphorylation catalyzed by SWE1.
CC   -!- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
CC       tyrosine + phosphate.
CC   -!- SIMILARITY: BELONGS TO THE MPI PHOSPHATASE FAMILY.
CC   -!- SIMILARITY: Contains 1 rhodanese domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; JC4846; AAA57223.1; -.
DR   EMBL; Z48502; CAA88402.1; -.
DR   EMBL; Z49213; CAA89152.1; -.
DR   PIR; A32386; A32386.
DR   HSSP; P30305; 1QB0.
DR   SGD; S0004639; MIH1.
DR   GO; GO:0000079; P:regulation of CDK activity; IGI.
DR   InterPro; IPR000751; MPI_Phosphatase.
DR   InterPro; IPR001763; Rhodanese-like.
DR   Pfam; PF00581; Rhodanese; 1.
DR   PRINTS; PR00716; MPIPHPHTASE.
DR   SMART; SM00450; RHOD; 1.
DR   PROSITE; PS50206; RHODANESE_3; 1.
KW   Cell division; Cell cycle; Mitosis; Hydrolase.
FT   DOMAIN       261     373       RHODANESE.
FT   ACT_SITE     320     320       BY SIMILARITY.
FT   CONFLICT     238     239       ND -> MT (IN REF. 1).
FT   CONFLICT     241     248       FPRTSPET -> SLEFLQKR (IN REF. 1).
FT   CONFLICT     365     367       FDN -> LT (IN REF. 1).
FT   CONFLICT     475     554       ILSKSSMSSNSNLSTSHMLLMDGLDIPSYFSFEDERGNHQQ
FT                                  VSGDEEQQGDFIFVGSDREDLPRPARRSLFPSLETEDKK
FT                                  -> S (IN REF. 1).
SQ   SEQUENCE   554 AA;  63357 MW;  DBD85D02CF48CA6F CRC64;

  Query Match             54.5%;  Score 36;  DB 1;  Length 554;
  Best Local Similarity   63.6%;  Pred. No. 25;
  Matches    7;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 SQRFSKIASNT 11
              |||:|| |  ||
Db        188 SQKFSNITQNT 198


RESULT 7
APCE_PORPU
ID   APCE_PORPU     STANDARD;      PRT;   886 AA.
AC   P51263;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   Phycobilisome linker polypeptide (Anchor polypeptide) (PBS-anchor
DE   protein).
GN   APCE.
OS   Porphyra purpurea.
OG   Chloroplast.
OC   Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae; Porphyra.
OX   NCBI_TaxID=2787;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Avonport;
RA   Reith M.E., Munholland J.;
RT   "Complete nucleotide sequence of the Porphyra purpurea chloroplast
RT   genome.";
RL   Plant Mol. Biol. Rep. 13:333-335(1995).
CC   -!- FUNCTION: THIS PROTEIN IS POSTULATED TO ACT BOTH AS TERMINAL
CC       ENERGY ACCEPTOR AND AS A LINKER POLYPEPTIDE THAT STABILIZES
CC       THE PHYCOBILISOME ARCHITECTURE.
CC   -!- SIMILARITY: OF THE REPEATED DOMAINS TO N-TERMINAL REGIONS OF
CC       PHYCOCYANIN ROD LINKER POLYPEPTIDES.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U38804; AAC08149.1; -.
DR   PIR; S73184; S73184.
DR   HSSP; P00315; 1B33.
DR   InterPro; IPR001297; PBS_linker_poly.
DR   InterPro; IPR001659; Phycobilisome.
DR   Pfam; PF00427; PBS_linker_poly; 3.
DR   Pfam; PF00502; Phycobilisome; 1.
DR   ProDom; PD000852; PBS_linker_poly; 3.
DR   ProDom; PD000340; Phycobilisome; 2.
KW   Phycobilisome; Electron transport; Photosynthesis; Repeat;
KW   Chloroplast.
SQ   SEQUENCE   886 AA;  100160 MW;  B21074C6D58C6203 CRC64;

  Query Match             54.5%;  Score 36;  DB 1;  Length 886;
```

```
  Best Local Similarity   53.8%;  Pred. No. 42;
  Matches    7;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 QRFSKIASNTQSR 14
              .| |: ': :|
Db        669 QRFKKVISSKSSR 681


RESULT 8
OSTP_RABIT
ID   OSTP_RABIT     STANDARD;      PRT:   311 AA.
AC   P31097; P46631;
DT   01-JUL-1993 (Rel. 25, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Osteopontin precursor (Bone sialoprotein 1) (Secreted phosphoprotein
DE   1) (SPP-1) (OC-1).
GN   SPP1.
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Osteoclast;
RX   MEDLINE=92360044; PubMed=1379809;
RA   Tezuka K.-I., Sato T., Kamioka H., Nijweide P.J., Tanaka K.,
RA   Matsuo T., Ohta M., Kurihara N., Hakeda Y., Kumegawa M.;
RT   "Identification of osteopontin in isolated rabbit osteoclasts.";
RL   Biochem. Biophys. Res. Commun. 186:911-917(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Myeloid;
RA   Nasu K., Ishida T., Hijiya N., Setoguchi M., Akizuki S.,
RA   Higuchi Y., Yamamoto S.;
RL   Submitted (JUN-1993) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Binds tightly to hydroxyapatite. Appears to form an
CC       integral part of the mineralized matrix. Probably important to
CC       cell-matrix interaction.
CC   -!- FUNCTION: Acts as a cytokine involved in enhancing production of
CC       interferon-gamma and interleukin-12 and reducing production of
CC       interleukin-10 and is essential in the pathway that leads to type
CC       1 immunity (By similarity).
CC   -!- SUBUNIT: LIGAND FOR INTEGRIN ALPHA-V/BETA-3.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- PTM: EXTENSIVELY PHOSPHORYLATED ON SERINE RESIDUES.
CC   -!- PTM: N- AND O-GLYCOSYLATED.
CC   -!- SIMILARITY: BELONGS TO THE OSTEOPONTIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D11411; BAA01993.1; -.
DR   EMBL; D16544; BAA03980.1; -.
DR   PIR; JC1191; JC1191.
DR   InterPro; IPR002038; Osteopontin.
DR   Pfam; PF00865; Osteopontin; 1.
DR   PRINTS; PR00216; OSTEOPONTIN.
DR   SMART; SM00017; OSTEO; 1.
DR   PROSITE; PS00884; OSTEOPONTIN; 1.
KW   Cytokine; Glycoprotein; Sialic acid; Biomineralization; Cell adhesion;
KW   Phosphorylation; Signal.
FT   SIGNAL       1     16       BY SIMILARITY.
FT   CHAIN       17    311       OSTEOPONTIN.
FT   SITE       155    157       CELL ATTACHMENT SITE.
FT   CARBOHYD   103    103       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CONFLICT    85     85       M -> V (IN REF. 2).
FT   CONFLICT   100    100       R -> Q (IN REF. 2).
FT   CONFLICT   128    128       E -> ESDE (IN REF. 2).
FT   CONFLICT   151    151       T -> I (IN REF. 2).
SQ   SEQUENCE   311 AA;  35172 MW;  69275CF63B7CB5A5 CRC64;

  Query Match             53.0%;  Score 35;  DB 1;  Length 311;
  Best Local Similarity   50.0%;  Pred. No. 21;
  Matches    7;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 SQRFSKIASNTQSR 14
              :'  ||::  :||:
Db        260 SQESSKVSQESQSR 273


RESULT 9
BEM4_YEAST
ID   BEM4_YEAST     STANDARD;      PRT:   633 AA.
AC   P39021;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   BEM4 protein.
GN   BEM4 OR YPL161C OR P2561.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96315665; PubMed=8754839;
RA   Mack D., Nishimura K., Dennehey B.K., Arbogast T., Parkinson J.,
RA   Toh-E A., Pringle J.R., Bender A., Matsui Y.;
RT   "Identification of the bud emergence gene BEM4 and its interactions
RT   with rho-type GTPases in Saccharomyces cerevisiae.";
RL   Mol. Cell. Biol. 16:4387-4395(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288c / AB972;
RX   MEDLINE=97103777; PubMed=8948103;
RA   Purnelle B., Coster F., Goffeau A.;
RT   "The sequence of 55 kb on the left arm of yeast chromosome XVI
RT   identifies a small nuclear RNA, a new putative protein kinase and two
RT   new putative regulators.";
RL   Yeast 12:1483-1492(1996).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288c / AB972;
RX   MEDLINE=97313271; PubMed=9169875;
RA   Bussey H., Storms R.K., Ahmed A., Albermann K., Allen E., Ansorge W.,
RA   Araujo R., Aparicio A., Barrell B.G., Badcock K., Benes V.,
RA   Botstein D., Bowman S., Brueckner M., Carpenter J., Cherry J.M.,
RA   Chung E., Churcher C.M., Coster F., Davis K., Davis R.W.,
RA   Dietrich F.S., Delius H., DiPaolo T., Dubois E., Duesterhoeft A.,
RA   Duncan M., Floeth M., Fortin N., Friesen J.D., Fritz C., Goffeau A.,
RA   Hall J., Hebling U., Heumann K., Hilbert H., Hillier L.,
RA   Hunicke-Smith S., Hyman R., Johnston M., Kalman S., Kleine K.,
RA   Komp C., Kurdi O., Lashkari D., Lew H., Lin A., Lin D., Louis E.J.,
RA   Marathe R., Messenguy F., Mewes H.-W., Mirtipati S., Moestl D.,
RA   Mueller-Auer S., Namath A., Nentwich U., Oefner P., Pearson D.,
RA   Petel F.X., Pohl T.M., Purnelle B., Rajandream M.A., Rechmann S.,
RA   Rieger M., Riles L., Roberts A.B., Rose C., Scharfe M., Scharfe M.,
RA   Scherens B., Schramm S., Schroeder M., Sdicu A.M., Settelin H.,
RA   Urrestarazu L.A., Ushinsky S., Vierendeels F., Vissers S., Voss H.,
RA   Walsh S.V., Wambutt R., Wang Y., Wedler E., Wedler H., Winnett E.,
RA   Zhong W.W., Zollner A., Vo D.H., Hani J.;
RT   "The nucleotide sequence of Saccharomyces cerevisiae chromosome XVI.";
RL   Nature 387:103-105(1997).
CC   -!- FUNCTION: PROBABLY ACTS AS A GTPASE-ACTIVATING PROTEIN (GAP) FOR
CC       CDC42, RHO1, RHO2, AND RHO4. NEGATIVE REGULATOR OF THE PHEROMONE-
CC       RESPONSE PATHWAY THROUGH THE STE20 PROTEIN KINASE. INVOLVED IN THE
CC       CONTROL OF POLARIZED CELL GROWTH.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
```

```
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed.  Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L27816; AAB47774.1; -.
DR   EMBL; X96770; CAA65560.1; -.
DR   EMBL; Z73517; CAA97866.1; -.
DR   PIR; S45177; S45177.
DR   SGD; S0006082; BEM4.
DR   GO; GO:0000283; P:establishment of cell polarity (sensu Sacch. ...; IMP.
DR   GO; GO:0007266; P:Rho protein signal transduction; IGI.
KW   GTPase activation; Pheromone response.
SQ   SEQUENCE   633 AA;  70992 MW;  22171643922650AC CRC64;

  Query Match             53.0%;  Score 35;  DB 1;  Length 633;
  Best Local Similarity   50.0%;  Pred. No. 46;
  Matches    7;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 SQRFSKIASNTQSR 14
              ||   ||:|| ':
Db        221 SQYIKKISSNVQNK 234


RESULT 10
SECA_MYCPN
ID   SECA_MYCPN     STANDARD;      PRT:   808 AA.
AC   P75559;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Preprotein translocase secA subunit.
GN   SECA OR MPN210 OR MP621.
OS   Mycoplasma pneumoniae.
OC   Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX   NCBI_TaxID=2104;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 29342 / M129;
RX   MEDLINE=97105885; PubMed=8948633;
RA   Himmelreich R., Hilbert H., Plagens H., Pirkl E., Li B.-C.,
RA   Herrmann R.;
RT   "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT   pneumoniae.";
RL   Nucleic Acids Res. 24:4420-4449(1996).
CC   -!- FUNCTION: INVOLVED IN PROTEIN EXPORT. INTERACTS WITH THE SECY/SECE
CC       SUBUNITS. SECA HAS A CENTRAL ROLE IN COUPLING THE HYDROLYSIS OF
CC       ATP TO THE TRANSFER OF PRE-SECRETORY PERIPLASMIC AND OUTER
CC       MEMBRANE PROTEINS ACROSS THE MEMBRANE (BY SIMILARITY).
CC   -!- SUBUNIT: PART OF THE PROKARYOTIC PROTEIN TRANSLOCATION APPARATUS
CC       WHICH COMPRISE SECA, SECB, SECD, SECE, SECF, SECG AND SECY
CC       (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC SIDE OF MEMBRANE
CC       (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE SECA FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed.  Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE000059; AAB96269.1; -.
DR   PIR; S73947; S73947.
DR   InterPro; IPR001650; Helicase_C.
DR   InterPro; IPR000185; SecA.
DR   Pfam; PF00271; Helicase_C; 1.
DR   Pfam; PF01043; SecA_protein; 1.
DR   PRINTS; PR00906; SECA.
DR   TIGRFAMs; TIGR00963; secA; 1.
DR   PROSITE; PS01312; SECA; 1.
KW   Protein transport; ATP-binding; Membrane; Translocation; Transport;
KW   Complete proteome.
FT   NP_BIND     102    109       ATP (POTENTIAL).
SQ   SEQUENCE   808 AA;  91810 MW;  B768F058015F892F CRC64;

  Query Match             53.0%;  Score 35;  DB 1;  Length 808;
  Best Local Similarity   57.1%;  Pred. No. 60;
  Matches    8;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 SQRFSKIASNTQSR 14
              |  |||: | || |
Db        569 SSFFSKMLSRTQKR 582


RESULT 11
MNN9_YEAST
ID   MNN9_YEAST     STANDARD;      PRT:   394 AA.
AC   P39107;
DT   01-FEB-1995 (Rel. 31, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Mannan polymerase complexes MNN9 subunit (M-pol 1 subunit MNN9) (M-
DE   Pol II subunit MNN9).
GN   MNN9 OR YPL050C.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=X2180;
RX   MEDLINE=94195815; PubMed=8146181;
RA   Yip C.L., Welch S.K., Klebl F., Gilbert T., Seidel P., Grant F.J.,
RA   O'Hara P.J., Mackay V.L.;
RT   "Cloning and analysis of the Saccharomyces cerevisiae MNN9 and MNN1
RT   genes required for complex glycosylation of secreted proteins.";
RL   Proc. Natl. Acad. Sci. U.S.A. 91:2723-2727(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288c / AB972;
RX   MEDLINE=97313271; PubMed=9169875;
RA   Bussey H., Storms R.K., Ahmed A., Albermann K., Allen E., Ansorge W.,
RA   Araujo R., Aparicio A., Barrell B.G., Badcock K., Benes V.,
RA   Botstein D., Bowman S., Brueckner M., Carpenter J., Cherry J.M.,
RA   Chung E., Churcher C.M., Coster F., Davis K., Davis R.W.,
RA   Dietrich F.S., Delius H., DiPaolo T., Dubois E., Duesterhoeft A.,
RA   Duncan M., Floeth M., Fortin N., Friesen J.D., Fritz C., Goffeau A.,
RA   Hall J., Hebling U., Heumann K., Hilbert H., Hillier L.,
RA   Hunicke-Smith S., Hyman R., Johnston M., Kalman S., Kleine K.,
RA   Komp C., Kurdi O., Lashkari D., Lew H., Lin A., Lin D., Louis E.J.,
RA   Marathe R., Messenguy F., Mewes H.-W., Mirtipati S., Moestl D.,
RA   Mueller-Auer S., Namath A., Nentwich U., Oefner P., Pearson D.,
RA   Petel F.X., Pohl T.M., Purnelle B., Rajandream M.A., Scharfe M.,
RA   Scherens B., Schramm S., Schroeder M., Sdicu A.M., Tettelin H.,
RA   Urrestarazu L.A., Ushinsky S., Vierendeels F., Vissers S., Voss H.,
RA   Walsh S.V., Wambutt R., Wang Y., Wedler E., Wedler H., Winnett E.,
RA   Zhong W.W., Zollner A., Vo D.H., Hani J.;
RT   "The nucleotide sequence of Saccharomyces cerevisiae chromosome XVI.";
RL   Nature 387:103-105(1997).
RN   [3]
RP   SEQUENCE OF 1-10, SUBCELLULAR LOCATION, AND SUBUNITS.
RX   MEDLINE=98096381; PubMed=9434768;
RA   Hashimoto H., Yoda K.;
RT   "Novel membrane protein complexes for protein glycosylation in the
RT   yeast Golgi apparatus.";
RL   Biochem. Biophys. Res. Commun. 241:682-686(1997).
RN   [4]
RP   SEQUENCE OF 142-152, AND SUBCELLULAR LOCATION.
RC   STRAIN=RSY455;
RX   MEDLINE=20558560; PubMed=11095735;
RA   Todorow Z., Spang A., Carmack E., Yates J., Schekman R.;
RT   "Active recycling of yeast Golgi mannosyltransferase complexes through
RT   the endoplasmic reticulum.";
```

```
RL   Proc. Natl. Acad. Sci. U.S.A. 97:13643-13648(2000).
RN   [5]
RP   ACTIVITY OF M-POL COMPLEXES, SUBUNITS, AND SUBCELLULAR LOCATION.
RX   MEDLINE=98094364; PubMed-9430634;
RA   Jungmann J., Munro S.;
RT   "Multi-protein complexes in the cis Golgi of Saccharomyces cerevisiae
RT   with alpha-1,6-mannosyltransferase activity.";
RL   EMBO J. 17:423-434(1998).
CC   -!- FUNCTION: The M-Pol I and M-Pol II complexes possess alpha-1,6-
CC       mannosyltransferase activity and are probably involved in the
CC       elongation of the mannan backbone of N-linked glycans on cell wall
CC       and periplasmic proteins.
CC   -!- PATHWAY: Glycosylation.
CC   -!- SUBUNIT: The M-Pol I complex contains MNN9 and VAN1. The M-Pol II
CC       complex is composed of ANP1, MNN9, MNN10, MNN11 and HOC1.
CC   -!- SUBCELLULAR LOCATION: Type II membrane protein. Cis-Golgi.
CC       Recycles between endoplasmic reticulum and Golgi.
CC   -!- SIMILARITY: BELONGS TO THE ANP1 / MNN9 / VAN1 FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L23752; AAA53677.1; -.
DR   EMBL; U44030; AAB68171.1; -.
DR   PIR; S43746; S43746.
DR   SGD; S0005971; MNN9.
DR   GO; GO:0000136; C:mannosyltransferase complex; IPI.
DR   GO; GO:0000009; F:alpha-1,6-mannosyltransferase activity; IDA.
DR   GO; GO:0000032; P:cell wall mannoprotein biosynthesis; TAS.
DR   GO; GO:0006487; P:N-linked glycosylation; IDA.
DR   InterPro; IPR005109; Anp1.
DR   Pfam; PF03452; Anp1; 1.
KW   Glycoprotein; Transmembrane; Signal-anchor; Golgi stack;
KW   Endoplasmic reticulum.
FT   INIT_MET      0      0
FT   DOMAIN        1     16       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     17     32       SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT                                (POTENTIAL).
FT   DOMAIN       33    394       LUMENAL (POTENTIAL).
SQ   SEQUENCE   394 AA;  45823 MW;  A6AA1F8EC8082F6A CRC64;

  Query Match             52.3%;  Score 34.5;  DB 1;  Length 394;
  Best Local Similarity   56.2%;  Pred. No. 35;
  Matches    9;  Conservative    2;  Mismatches    2;  Indels    3;  Gaps    1;

Qy          1 SQRFSKIA---SNTQS 13
              :|||||!     :::|
Db        175 TQRFSKITILRQNSQS 190


RESULT 12
SLAP_LACHE
ID   SLAP_LACHE     STANDARD;      PRT;   439 AA.
AC   P38059;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   S-layer protein precursor (Surface layer protein).
GN   SLPH.
OS   Lactobacillus helveticus.
OC   Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC   Lactobacillus.
OX   NCBI_TaxID=1587;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CNRZ 892, and CNRZ 1269;
RA   Callegari M.L., Cocconcelli P.S., Morelli L.;
RL   Submitted (NOV-1995) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE OF 31-53.
RC   STRAIN=ATCC 12046;
RA   Lortal S., van Heijenoort J., Gruber K., Sleytr U.B.;
RT   "S-layer of Lactobacillus helveticus ATCC 12046: isolation, chemical
RT   characterization and re-formation after extraction with lithium
RT   chloride.";
RL   J. Gen. Microbiol. 138:611-618(1992).
CC   -!- FUNCTION: THE S-LAYER IS A PARACRYSTALLINE MONO-LAYERED ASSEMBLY
CC       OF PROTEINS WHICH COAT THE SURFACE OF BACTERIA.
CC   -!- SUBCELLULAR LOCATION: CELL WALL. THIS BACTERIUM IS COVERED BY A
CC       S-LAYER WITH HEXAGONAL SYMMETRY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X91199; CAA62606.1; -.
DR   EMBL; X92752; CAA63409.1; -.
DR   InterPro; IPR004993; SLAP.
DR   Pfam; PF03217; SLAP; 1.
KW   Glycoprotein; Cell wall; S-layer; Signal.
FT   SIGNAL        1     30
FT   CHAIN        31    439       S-LAYER PROTEIN.
FT   VARIANT     242    242       Q -> P (IN STRAIN CNRZ 1269).
SQ   SEQUENCE   439 AA;  46638 MW;  54D7DD98B8113AB2 CRC64;

  Query Match             51.5%;  Score 34;  DB 1;  Length 439;
  Best Local Similarity   54.5%;  Pred. No. 49;
  Matches    6;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 QRFSKIASNTQ 12
              |!: || :|!:
Db        419 QRYYKIGANTE 429


RESULT 13
AATM_LUPAN
ID   AATM_LUPAN     STANDARD;      PRT;   454 AA.
AC   P26563;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Aspartate aminotransferase-P2, mitochondrial precursor (EC 2.6.1.1)
DE   (Transaminase A) (Fragment).
OS   Lupinus angustifolius (Narrow-leaved blue lupine).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids I; Fabales; Fabaceae; Papilionoideae; Genisteae; Lupinus.
OX   NCBI_TaxID=3871;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Uniharvest; TISSUE=Root nodules;
RX   MEDLINE=92322978; PubMed=1623192;
RA   Reynolds P.H.S., Smith L.A., Jones W.T., Dickson J.M.J., Jones S.J.,
RA   Rodber K., Liddane C.P.;
RT   "Molecular cloning of a cDNA encoding aspartate aminotransferase-P2
RT   from lupin root nodules.";
RL   Plant Mol. Biol. 19:465-472(1992).
CC   -!- FUNCTION: IMPORTANT FOR THE METABOLISM OF AMINO ACIDS AND KREBS-
CC       CYCLE RELATED ORGANIC ACIDS. IN PLANTS, IT IS INVOLVED IN NITROGEN
CC       METABOLISM AND IN ASPECTS OF CARBON AND ENERGY METABOLISM.
CC   -!- CATALYTIC ACTIVITY: L-aspartate + 2-oxoglutarate = oxaloacetate +
CC       L-glutamate.
CC   -!- COFACTOR: Pyridoxal phosphate.
CC   -!- SUBUNIT: Homodimer (By similarity).
CC   -!- SUBCELLULAR LOCATION: Mitochondrial matrix.
CC   -!- MISCELLANEOUS: In eukaryotes there are cytoplasmic, mitochondrial
CC       and chloroplastic isozymes.
```

```
CC   -!- SIMILARITY: BELONGS TO CLASS-I OF PYRIDOXAL-PHOSPHATE-DEPENDENT
CC       AMINOTRANSFERASES.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X59761; CAA42430.1; -.
DR   PIR; S22465; XNYLB.
DR   HSSP; P00508; 7AAT.
DR   InterPro; IPR004839; Aminotransfl/2.
DR   InterPro; IPR000796; Asptransf_sub.
DR   InterPro; IPR004838; NHtransf_1.
DR   Pfam; PF00155; aminotran_1_2; 1.
DR   PRINTS; PR00799; TRANSAMINASE.
DR   PROSITE; PS00105; AA_TRANSFER_CLASS_1; 1.
KW   Transferase; Aminotransferase; Pyridoxal phosphate; Mitochondrion;
KW   Transit peptide.
FT   NON_TER       1      1
FT   TRANSIT       <1     49        MITOCHONDRION.
FT   CHAIN         50     454       ASPARTATE AMINOTRANSFERASE P2.
FT   BINDING      299     299       PYRIDOXAL PHOSPHATE (BY SIMILARITY).
SQ   SEQUENCE   454 AA;  49915 MW;  744CC12F1D81E739 CRC64;

  Query Match             51.5%;  Score 34;  DB 1;  Length 454;
  Best Local Similarity   69.2%;  Pred. No. 51;
  Matches    9;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 SQRFSKIASNTQS 13
              | |||:  ||. |
Db         23 SLRFSKEQSNTFS 35


RESULT 14
FLII_BUCAI
ID   FLII_BUCAI     STANDARD;      PRT:   467 AA.
AC   P57178;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Flagellum-specific ATP synthase (EC 3.6.3.14).
GN   FLII OR BU076.
OS   Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
OS   symbiotic bacterium).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Buchnera.
OX   NCBI_TaxID=118099;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Tokyo 1998;
RX   MEDLINE=20445173; PubMed=10993077;
RA   Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;
RT   "Genome sequence of the endocellular bacterial symbiont of aphids
RT   Buchnera sp. APS.";
RL   Nature 407:81-86(2000).
CC   -!- FUNCTION: PROBABLE CATALYTIC SUBUNIT OF A PROTEIN TRANSLOCASE FOR
CC       FLAGELLUM-SPECIFIC EXPORT, OR A PROTON TRANSLOCASE INVOLVED IN
CC       LOCAL CIRCUITS AT THE FLAGELLUM. MAY BE INVOLVED IN A SPECIALIZED
CC       PROTEIN EXPORT PATHWAY THAT PROCEEDS WITHOUT SIGNAL PEPTIDE
CC       CLEAVAGE (BY SIMILARITY).
CC   -!- CATALYTIC ACTIVITY: ATP + H(2)O + H(+)(In) = ADP + phosphate +
CC       H(+)(Out).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC   -!- SIMILARITY: Belongs to the ATPase alpha/beta chains family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AP001118; BAB12796.1; -.
DR   InterPro; IPR003593; AAA_ATPase.
DR   InterPro; IPR000194; ATPase_a/bcentre.
DR   InterPro; IPR005714; FliI_YscN.
DR   Pfam; PF00006; ATP-synt_ab; 1.
DR   SMART; SM00382; AAA; 1.
DR   TIGRFAMs; TIGR01026; fliI_yscN; 1.
DR   PROSITE; PS00152; ATPASE_ALPHA_BETA; 1.
KW   Hydrolase; Hydrogen ion transport; ATP synthesis; ATP-binding;
KW   Transport; Protein transport; Flagella; Complete proteome.
FT   NP_BIND     195    202       ATP (BY SIMILARITY).
SQ   SEQUENCE   467 AA;  51946 MW;  0C87A34CDC29A732 CRC64;

  Query Match             51.5%;  Score 34;  DB 1;  Length 467;
  Best Local Similarity   87.5%;  Pred. No. 52;
  Matches    7;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 FSKIASNT 11
              ||:|:||:
Db        109 FSKIASNS 116


RESULT 15
NU5C_HORVU
ID   NU5C_HORVU     STANDARD;      PRT:   705 AA.
AC   Q32440;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   NAD(P)H-quinone oxidoreductase chain 5, chloroplast (EC 1.6.5.-)
DE   (NAD(P)H dehydrogenase, chain 5) (NADH-plastoquinone oxidoreductase
DE   chain 5) (Fragment).
GN   NDHF.
OS   Hordeum vulgare (Barley).
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC   Triticeae; Hordeum.
OX   NCBI_TaxID=4513;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Leaf;
RA   Clark L.G., Zhang W., Wendel J.F.;
RL   Submitted (APR-1995) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: NAD(P)H + plastoquinone = NAD(P)(+) +
CC       plastoquinol.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U22003; AAA64207.1; -.
DR   PIR; T04400; T04400.
DR   InterPro; IPR003916; NADHub_oxred5.
DR   InterPro; IPR001750; Oxidored_q1.
DR   InterPro; IPR002128; Oxidored_q1_C.
DR   InterPro; IPR001516; Oxidored_q1_N.
DR   Pfam; PF00361; oxidored_q1; 1.
DR   Pfam; PF01010; oxidored_q1_C; 1.
DR   Pfam; PF00662; oxidored_q1_N; 1.
DR   PRINTS; PR01434; NADHDHGNASE5.
KW   Oxidoreductase; NAD; NADP; Quinone; Plastoquinone; Chloroplast.
FT   NON_TER       1      1
FT   NON_TER     705    705
SQ   SEQUENCE   705 AA;  78841 MW;  ACC63E07C070ADE5 CRC64;
```

Query Match 51.5%; Score 34; DB 1; Length 705;
Best Local Similarity 50.0%; Pred. No. 81;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

```
Qy          1 SQRFSKIASNFQSR 14
              ::   ||  ||:::
Db        496 SQNIPKIQGNTRNR 509
```

Search completed: September 24, 2003, 13:42:36
Job time : 9.80859 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 24, 2003, 13:33:45 ; Search time 45.4607 Seconds
(without alignments)
79.469 Million cell updates/sec

Title: US-09-744-282-8
Perfect score: 66
Sequence: 1 SQRFSKIASNTQSR 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	66	100.0	383	4	Q96HI7	Q96hi7 homo sapien
2	66	100.0	383	4	Q8IVA4	Q8iva4 homo sapien
3	66	100.0	383	5	Q29411	Q29411 sus scrofa
4	60	90.9	382	5	O18949	O18949 bos taurus
5	53	80.3	389	6	Q8SPQ0	Q8spq0 capra hircu
6	47	71.2	352	11	Q9WTV1	Q9wtv1 rattus norv
7	46	69.7	381	11	Q99J84	Q99j84 mus musculu
8	46	69.7	389	11	Q8BKL8	Q8bkl8 mus musculu
9	41	62.1	705	8	Q9T2A6	Q9t2a6 puella olyr
10	41	62.1	787	16	Q8DSC6	Q8dsc6 streptococc
11	40	60.6	454	5	Q9TMV4	Q9tmv4 stipa barba
12	39	59.1	700	8	Q31958	Q31958 chusquea ci
13	39	59.1	702	8	Q32173	Q32173 ehrharta ca
14	39	59.1	703	8	Q32532	Q32532 lithachne p
15	39	59.1	704	8	Q9T2A8	Q9t2a8 sacrea macu
16	39	59.1	704	8	Q32740	Q32740 olyra latif
17	39	59.1	705	8	Q32142	Q32142 diarrhena o
18	39	59.1	706	8	Q32522	Q32522 lithachne h
19	39	59.1	929	5	Q8MY79	Q8my79 haemaphysal
20	38	57.6	84	8	P92321	P92321 nardus stri
21	38	57.6	212	8	P92313	P92313 lygeum spar
22	38	57.6	212	8	P92350	P92350 stipa parvi
23	38	57.6	520	10	Q9SU19	Q9su19 arabidopsis
24	38	57.6	694	8	Q8WJZ6	Q8wjz6 anisopogon
25	38	57.6	695	8	Q8WJZ5	Q8wjz5 ampelodesmo
26	38	57.6	695	8	Q8WJZ4	Q8wjz4 stipa barba
27	38	57.6	695	8	Q8WJY8	Q8wjy8 stipagrosti
28	38	57.6	697	8	Q8WJZ8	Q8wjz8 lygeum spar
29	38	57.6	698	8	Q8WJZ3	Q8wjz3 nassella vi
30	38	57.6	700	8	Q31863	Q31863 bambusa aff
31	38	57.6	701	8	Q32172	Q32172 oragrostis
32	38	57.6	701	8	Q9T2B0	Q9t2b0 buergersioc
33	38	57.6	701	8	Q9T2A4	Q9t2a4 alvimia gra
34	38	57.6	702	8	Q32334	Q32334 guadua pani
35	38	57.6	702	8	Q9T296	Q9t296 neurolepis
36	38	57.6	702	8	Q32589	Q32589 micraira la
37	38	57.6	702	8	Q9T2A3	Q9t2a3 melocanna b
38	38	57.6	702	8	Q9T2A2	Q9t2a2 rhipidoclad
39	38	57.6	702	8	Q9T2A7	Q9t2a7 sasa varieg
40	38	57.6	702	8	Q9T299	Q9t299 nastus elat
41	38	57.6	702	8	Q33053	Q33053 streptogyna
42	38	57.6	702	8	Q31926	Q31926 bambusa ste
43	38	57.6	702	8	Q9T2A1	Q9t2a1 otatea acum
44	38	57.6	703	8	Q31800	Q31800 arundinaria
45	38	57.6	703	8	Q31878	Q31878 brachyelytr

ALIGNMENTS

RESULT 1
Q96HI7
ID Q96HI7 PRELIMINARY; PRT; 383 AA.
AC Q96HI7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Similar to chitinase 3-like 1 (Cartilage glycoprotein-39).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC008568; AAH08568.1; -.
DR InterPro; IPR001223; Glyco_hydro_18.
DR InterPro; IPR005829; Sug_transporter.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR SMART; SM00636; Glyco_18; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
SQ SEQUENCE 383 AA; 42625 MW; 6C9EF133BDC7C2D1 CRC64;

Query Match 100.0%; Score 66; DB 4; Length 383;
Best Local Similarity 100.0%; Pred. No. 0.00019;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SQRFSKIASNTQSR 14
||||||||||||||
Db 103 SQRFSKIASNTQSR 116

RESULT 2
Q8IVA4
ID Q8IVA4 PRELIMINARY; PRT; 383 AA.
AC Q8IVA4;

```
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Similar to chitinase 3-like 1 (Cartilage glycoprotein-39).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RA   Strausberg R.;
RL   Submitted (OCT-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC038354; AAH38354.1; -.
SQ   SEQUENCE   383 AA;  42526 MW;  76345433BDC7C6C0 CRC64;

  Query Match             100.0%;  Score 66;  DB 4;  Length 383;
  Best Local Similarity   100.0%;  Pred. No. 0.00019;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SQRFSKIASNTQSR 14
              ||||||||||||||
Db        103 SQRFSKIASNTQSR 116


RESULT 3
Q29411
ID   Q29411      PRELIMINARY;      PRT;   383 AA.
AC   Q29411;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   38 kDa heparin-binding glycoprotein.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Smooth muscle;
RX   MEDLINE=95286589; PubMed=7768902;
RA   Shackelton L.M., Mann D.M., Millis A.J.;
RT   "Identification of a 38-kDa heparin-binding glycoprotein (gp38k) in
RT   differentiating vascular smooth muscle cells as a member of a group of
RT   proteins associated with tissue remodeling.";
RL   J. Biol. Chem. 270:13076-13083(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Smooth muscle;
RA   Shackelton L.M., Mann D.M., Millis A.J.F.;
RL   Submitted (JAN-1995) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U19900; AAA86482.1; -.
DR   EMBL; Z47803; CAA87764.1; -.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   ProDom; PD000471; Glyco_hydro_18; 1.
DR   SMART; SM00636; Glyco_18; 1.
KW   Heparin-binding.
SQ   SEQUENCE   383 AA;  42443 MW;  3D1039F4391C0BDC4 CRC64;

  Query Match             100.0%;  Score 66;  DB 6;  Length 383;
  Best Local Similarity   100.0%;  Pred. No. 0.00019;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SQRFSKIASNTQSR 14
              ||||||||||||||
Db        103 SQRFSKIASNIQSR 116


RESULT 4
O18949
ID   O18949      PRELIMINARY;      PRT;   332 AA.
AC   O18949;
```

```
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Chitinase-like protein 1 (Fragment).
GN   CLP-1.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Recklies A.D., White C.;
RT   "Expression of chitinase-like protein 1 (CLP-1) in bovine
RT   chondrocytes.";
RL   Submitted (JUN-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF011373; AAB64304.1; -.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   ProDom; PD000471; Glyco_hydro_18; 1.
DR   SMART; SM00636; Glyco_18; 1.
FT   NON_TER       1       1
FT   NON_TER     332     332
SQ   SEQUENCE   332 AA;  37434 MW;  03F163323486408F CRC64;

  Query Match             90.9%;  Score 60;  DB 6;  Length 332;
  Best Local Similarity   92.9%;  Pred. No. 0.0024;
  Matches   13;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 SQRFSKIASNTQSR 14
              |||||||||| ||||
Db         79 SQRFSKIASKTQSR 92


RESULT 5
Q8SPQ0
ID   Q8SPQ0      PRELIMINARY;      PRT;   383 AA.
AC   Q8SPQ0;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   BP40 precursor.
OS   Capra hircus (Goat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Caprinae; Capra.
OX   NCBI_TaxID=9925;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Mammary gland;
RA   Paramasivam M., Saravanan K., Sharma S., Mohanty A.K., Yadav S.,
RA   Srinivasan A., Singh T.P.;
RT   "Goat mammary gland mRNA for a novel protein, BP40, complete coding
RT   region.";
RL   Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY081150; AAL87007.1; -.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   ProDom; PD000471; Glyco_hydro_18; 1.
DR   SMART; SM00636; Glyco_18; 1.
KW   Signal.
FT   SIGNAL        1      16       POTENTIAL.
SQ   SEQUENCE   383 AA;  42893 MW;  17655ED4BE4E9F5F CRC64;

  Query Match             80.3%;  Score 53;  DB 6;  Length 383;
  Best Local Similarity   84.6%;  Pred. No. 0.063;
  Matches   11;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 QRFSKIASNIQSR 14
              :|||||| ||||
Db        104 ERFSKIASKTQSR 116
```

```
RESULT 6
Q9WTV1
ID   Q9WTV1      PRELIMINARY;      PRT;   352 AA.
AC   Q9WTV1;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Glycoprotein-39 (Fragment).
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Lewis;
RA   Wendling C., Boots A.M.H., van Eden W.;
RT   "Cloning of the rat homologue of human cartilage glycoprotein-39 a
RT   potential autoantigen in arthritis.";
RL   Submitted (APR-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF062038; AAD22610.1; -.
DR   HSSP; P07254; 1CTN.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   ProDom; PD000471; Glyco_hydro_18; 1.
DR   SMART; SM00636; Glyco_18; 1.
FT   NON_TER       1      1
FT   NON_TER     352    352
SQ   SEQUENCE   352 AA;  39391 MW;  CBDE991810AC9360 CRC64;

  Query Match             71.2%;  Score 47;  DB 11;  Length 352;
  Best Local Similarity   64.3%;  Pred. No. 0.84;
  Matches    9;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 SQRFSKIASNTQSR 14
              |:|:|:| :| :||
Db         82 SERFSRIVSNAKSR 95


RESULT 7
Q99J84
ID   Q99J84      PRELIMINARY;      PRT;   381 AA.
AC   Q99J84;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Similar to chitinase 3-like 1 (Cartilage glycoprotein-39).
GN   CHI3L1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Strausberg R.;
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Strausberg R.;
RL   Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Strausberg R.;
RL   Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC005611; AAH05611.1; -.
DR   EMBL; BC003780; AAH03780.1; -.
DR   EMBL; BC004734; AAH04734.1; -.
DR   MGD; MGI:1340899; Chi3l1.
DR   InterPro; IPR001223; Glyco_hydro_18.
DR   Pfam; PF00704; Glyco_hydro_18; 1.
DR   ProDom; PD000471; Glyco_hydro_18; 1.
DR   SMART; SM00636; Glyco_18; 1.
SQ   SEQUENCE   381 AA;  42979 MW;  EF6588C5AE9D4453 CRC64;

  Query Match             69.7%;  Score 46;  DB 11;  Length 381;
  Best Local Similarity   69.2%;  Pred. No. 1.4;
  Matches    9;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 QRFSKIASNTQSR 14
              :|||:|:|||: !
Db        105 KRFSEIASNTERR 117


RESULT 8
Q8BKL8
ID   Q8BKL8      PRELIMINARY;      PRT;   389 AA.
AC   Q8BKL8;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Similar to chitinase 3-like 2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Dorsal root ganglion;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
DR   EMBL; AK051475; BAC34654.1; -.
SQ   SEQUENCE   389 AA;  43893 MW;  9E7D66069B233834 CRC64;

  Query Match             69.7%;  Score 46;  DB 11;  Length 389;
  Best Local Similarity   69.2%;  Pred. No. 1.5;
  Matches    9;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 QRFSKIASNTQSR 14
              :|||:|:|||: !
Db        113 KRFSEIASNTERR 125


RESULT 9
Q9T2A6
ID   Q9T2A6      PRELIMINARY;      PRT;   705 AA.
AC   Q9T2A6;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   NADH dehydrogenase (Fragment).
GN   NDHF.
OS   Puelia olyriformis.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Puelioideae; Puelia.
OX   NCBI_TaxID=93677;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Zhang W., Clark L.G.;
RT   "Phylogeny and classification of the Bambusoideae (Poaceae).";
RL   (In) Wilson K.L., Morrison D.A. (eds.);
RL   Monocots: systematics and evolution, pp.1-1, CSIRO, Melbourne (1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Zhang W., Clark L.G.;
RL   Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF192345; AAF25173.1; -.
DR   InterPro; IPR003916; NADHub_oxred5.
DR   InterPro; IPR001750; Oxidored_q1.
DR   InterPro; IPR002128; Oxidored_q1_C.
DR   InterPro; IPR001516; Oxidored_q1_N.
DR   Pfam; PF00361; oxidored_q1; 1.
```

```
DR   Pfam; PF01010; oxidored_q1_C; 1.
DR   Pfam; PF00662; oxidored_q1_N; 1.
DR   PRINTS; PR01434; NADHDHGNASE5.
KW   NAD; Oxidoreductase; Plastoquinone; Chloroplast.
FT   NON_TER       1       1
FT   NON_TER     705     705
SQ   SEQUENCE   705 AA;  78805 MW;  D43104AB46266B38 CRC64;

  Query Match           62.1%;  Score 41;  DB 8;  Length 705;
  Best Local Similarity   64.3%;  Pred. No. 26;
  Matches    9;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 SQRFSKIASNTQSR 14
              |:  |||  |:|:
Db        496 SQNISKIHGNTRSR 509


RESULT 10
Q8DSC6
ID   Q8DSC6      PRELIMINARY;      PRT:   787 AA.
AC   Q8DSC6;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Putative exodeoxyribonuclease V, similar to conjugation transfer
DE   protein.
GN   SMU.1875.
OS   Streptococcus mutans.
OC   Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=1309;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=UA159 / ATCC 700610 / Serotype C;
RX   MEDLINE=22295063; PubMed=12397186;
RA   Ajdic D., McShan W.M., McLaughlin R.E., Savic G., Chang J.,
RA   Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y.,
RA   Li S., Zhu H., Najar F., Lai H., White J., Roe B.A., Ferretti J.J.;
RT   "Genome sequence of Streptococcus mutans UA159, a cariogenic dental
RT   pathogen.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
DR   EMBL; AE015013; AAN59492.1; -.
KW   Complete proteome.
SQ   SEQUENCE   787 AA;  88886 MW;  CB4D67B560F91DDF CRC64;

  Query Match           62.1%;  Score 41;  DB 16;  Length 787;
  Best Local Similarity   75.0%;  Pred. No. 29;
  Matches    9;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 SQRFSKIASNTQ 12
              :| || ||||
Db        446 NQLFSAIASNIQ 457


RESULT 11
Q9TMV4
ID   Q9TMV4      PRELIMINARY;      PRT:   454 AA.
AC   Q9TMV4;
DT   01-MAY-2000 (TrEMBLrel. 14, Created)
DT   01-MAY-2000 (TrEMBLrel. 14, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   NADH dehydrogenase F (Fragment).
GN   NDHF.
OS   Stipa barbata.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC   Stipeae; Stipa.
OX   NCBI_TaxID=100941;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Clark L.G., Kobayashi M., Mathews S., Spangler R.E., Kellogg E.A.;
RT   "The Puelioideae, a new subfamily of Poaceae.";
RL   Syst. Bot. 0:0-0(1999).
DR   EMBL; AF164781; AAD50785.1; -.
DR   InterPro; IPR003916; NADHub_oxred5.
DR   InterPro; IPR001750; Oxidored_q1.
DR   InterPro; IPR002128; Oxidored_q1_C.
DR   InterPro; IPR001516; Oxidored_q1_N.
DR   Pfam; PF00361; oxidored_q1; 2.
DR   Pfam; PF01010; oxidored_q1_C; 1.
DR   Pfam; PF00662; oxidored_q1_N; 1.
DR   PRINTS; PR01434; NADHDHGNASE5.
KW   NAD; Oxidoreductase; Plastoquinone; Chloroplast.
FT   NON_TER       1       1
FT   NON_TER     454     454
SQ   SEQUENCE   454 AA;  50421 MW;  28F41CCAD8D3536E CRC64;

  Query Match           60.6%;  Score 40;  DB 8;  Length 454;
  Best Local Similarity   57.1%;  Pred. No. 25;
  Matches    8;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 SQRFSKIASNTQSR 14
              :|  |||  ||:|:|
Db        427 SQNISKIXXNTRNR 440


RESULT 12
Q31958
ID   Q31958      PRELIMINARY;      PRT:   700 AA.
AC   Q31958;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   NADH dehydrogenase F (Fragment).
GN   NDHF.
OS   Chusquea circinata.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
CC   Bambuscideae; Bambuseae; Chusquea.
OX   NCBI_TaxID=35710;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Leaf;
RA   Clark L.G., Zhang W., Wendel J.F.;
RT   "A Phylogeny of the Grass Family Based on ndhF Sequence Data.";
RL   Submitted (FEB-1995) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U21991; AAA64195.1; -.
DR   InterPro; IPR003916; NADHub_oxred5.
DR   InterPro; IPR001750; Oxidored_q1.
DR   InterPro; IPR002128; Oxidored_q1_C.
DR   InterPro; IPR001516; Oxidored_q1_N.
DR   Pfam; PF00361; oxidored_q1; 1.
DR   Pfam; PF01010; oxidored_q1_C; 1.
DR   Pfam; PF00662; oxidored_q1_N; 1.
DR   PRINTS; PR01434; NADHDHGNASE5.
KW   NAD; Oxidoreductase; Plastoquinone; Chloroplast.
FT   NON_TER       1       1
FT   NON_TER     700     700
SQ   SEQUENCE   700 AA;  78212 MW;  AD90A738F9BA3EDD CRC64;

  Query Match           59.1%;  Score 39;  DB 8;  Length 700;
  Best Local Similarity   57.1%;  Pred. No. 63;
  Matches    8;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 SQRFSKIASNTQSR 14
              :|  |||  ||:|:|
Db        493 SQNISKIPGNIRNR 506


RESULT 13
Q32173
ID   Q32173      PRELIMINARY;      PRT:   702 AA.
AC   Q32173;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
```

```
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   NADH dehydrogenase F (Fragment).
GN   NDHF.
OS   Ehrharta calycina (perennial veldt grass).
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Ehrharteae; Ehrharta.
OX   NCBI_TaxID=38695;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Leaf;
RA   Clark L.G., Zhang W., Wendel J.F.;
RT   "A Phylogeny of the Grass Family Based on ndhF Sequence Data.";
RL   Submitted (FEB-1995) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U21996; AAA64200.1; -.
DR   InterPro; IPR003916; NADHub_oxred5.
DR   InterPro; IPR001750; Oxidored_q1.
DR   InterPro; IPR002128; Oxidored_q1_C.
DR   InterPro; IPR001516; Oxidored_q1_N.
DR   Pfam; PF00361; oxidored_q1; 1.
DR   Pfam; PF01010; oxidored_q1_C; 1.
DR   Pfam; PF00662; oxidored_q1_N; 1.
DR   PRINTS; PR01434; NADHDHGNASE5.
KW   NAD; Oxidoreductase; Plastoquinone; Chloroplast.
FT   NON_TER       1       1
FT   NON_TER     702     702
SQ   SEQUENCE   702 AA;  78264 MW;  EB827ZBABCEDE2F9 CRC64;

  Query Match             59.1%;  Score 39;  DB 8;  Length 702;
  Best Local Similarity   57.1%;  Pred. No. 63;
  Matches    8;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 SQRFSKIASNTQSR 14
              |:  .|| '|:::|
Db        496 SKNISKIPSNTRNR 509


RESULT 14
O32532
ID   Q32532      PRELIMINARY;      PRT;   703 AA.
AC   Q32532;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   NADH dehydrogenase F (Fragment).
GN   NDHF.
OS   Lithachne pauciflora.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Bambusoideae; Olyreae; Lithachne.
OX   NCBI_TaxID=38736;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Leaf;
RA   Clark L.G., Zhang W., Wendel J.F.;
RT   "A Phylogeny of the Grass Family Based on ndhF Sequence Data.";
RL   Submitted (FEB-1995) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U21978; AAA64697.1; -.
DR   InterPro; IPR003916; NADHub_oxred5.
DR   InterPro; IPR001750; Oxidored_q1.
DR   InterPro; IPR002128; Oxidored_q1_C.
DR   InterPro; IPR001516; Oxidored_q1_N.
DR   Pfam; PF00361; oxidored_q1; 1.
DR   Pfam; PF01010; oxidored_q1_C; 1.
DR   Pfam; PF00662; oxidored_q1_N; 1.
DR   PRINTS; PR01434; NADHDHGNASE5.
KW   NAD; Oxidoreductase; Plastoquinone; Chloroplast.
FT   NON_TER       1       1
FT   NON_TER     703     703
SQ   SEQUENCE   703 AA;  78384 MW;  05DCBF5E823E26E8 CRC64;

  Query Match             59.1%;  Score 39;  DB 8;  Length 703;
  Best Local Similarity   57.1%;  Pred. No. 63;
  Matches    8;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 SQRFSKIASNTQSR 14
              ||  ||| |||::|
Db        495 SQNISKIKGNTRNR 508


RESULT 15
Q9T2A8
ID   Q9T2A8      PRELIMINARY;      PRT;   704 AA.
AC   Q9T2A8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   NADH dehydrogenase (Fragment).
GN   NDHF.
OS   Sucrea maculata.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Bambusoideae; Olyreae; Sucrea.
OX   NCBI_TaxID=47456;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Zhang W., Clark L.G.;
RT   "Phylogeny and classification of the Bambusoideae (Poaceae).";
RL   (In) Wilson K.L., Morrison D.A. (eds.);
RL   Monocots: systematics and evolution, pp.1-1, CSIRO, Melbourne (1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Zhang W., Clark L.G.;
RL   Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF182343; AAF25171.1; -.
DR   InterPro; IPR003916; NADHub_oxred5.
DR   InterPro; IPR001750; Oxidored_q1.
DR   InterPro; IPR002128; Oxidored_q1_C.
DR   InterPro; IPR001516; Oxidored_q1_N.
DR   Pfam; PF00361; oxidored_q1; 1.
DR   Pfam; PF01010; oxidored_q1_C; 1.
DR   Pfam; PF00662; oxidored_q1_N; 1.
DR   PRINTS; PR01434; NADHDHGNASE5.
KW   NAD; Oxidoreductase; Plastoquinone; Chloroplast.
FT   NON_TER       1       1
FT   NON_TER     704     704
SQ   SEQUENCE   704 AA;  79013 MW;  035AA18D95DE72B1 CRC64;

  Query Match             59.1%;  Score 39;  DB 8;  Length 704;
  Best Local Similarity   57.1%;  Pred. No. 63;
  Matches    8;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 SQRFSKIASNTQSR 14
              ||  ||| |||::|
Db        496 SQNISKIQGNTRNR 509


Search completed: September 24, 2003, 13:47:28
Job time : 46.4607 secs
```